Exhibit 8

CHAPTER 6

ECONOMETRICS AND REGRESSION ANALYSIS

Econometrics refers to a set of statistical methods or tools that economists use to assess antitrust damages. Regression analysis is the most common tool of econometrics. It enables one to uncover the relationship between a dependent variable (the outcome being modeled, such as prices) and one or more explanatory variables (the potential influences on the dependent variable, such as anticompetitive conduct, supply factors including costs or capacity, and demand factors).[1] Regressions can prove useful in answering the basic "counterfactual" question: What would market outcomes (e.g., prices) have been in the absence of the anticompetitive act? By their very nature, these but-for market outcomes are unobserved and, as such, must be estimated by the damages expert.

Market outcomes are often the result of a complex interaction among a large number of factors. For example, market prices were likely affected by a wide variety of demand and supply factors unrelated to the alleged anticompetitive act. Isolating and measuring the effect of an alleged anticompetitive act on price (or on other market outcomes) requires properly accounting for these other factors. Otherwise, one might attribute to the alleged anticompetitive act the effect of one or more of these other factors or miss the effects of the act in question by failing to control for a relevant factor.

Econometric models—particularly those involving regression analysis—are uniquely suited to isolating the effect of a single factor on the market factor of interest, while properly accounting for other relevant factors. The measured effect of an explanatory variable, controlling for other factors, is referred to as an estimate of the "partial effect" of the explanatory variables on the dependent variable. In other words, an explanatory variable's partial effect is the change in the dependent variable that would result from a change in that explanatory variable, holding all of the other explanatory variables constant.[2] Thus, when correctly implemented, econometric techniques can isolate and measure the effect of a single explanatory factor—such as the impact of the alleged

1. *See* Jeffrey M. Wooldridge, Econometric Analysis of Cross Section and Panel Data 3-4 (2d ed. 2010).
2. *See* William Greene, Econometric Analysis 36 (7th ed. 2010).

conduct—on the economic outcomes that are relevant when estimating damages.

The key role of econometrics in legal proceedings is to use the available data to provide accurate and reliable measures of the economic impact of the alleged conduct for the finders of fact (judges or juries). This chapter describes the legal requirements, the methods, and a wide range of issues in the use of econometrics and statistics in expert testimony. To begin, the damages methodology must be designed to causally determine the impact of the alleged conduct. Implementation must address issues of: data, model specification, estimation, interpretation of estimation results, and hypothesis testing. The following discussion explains the use of econometrics to implement before-during-after and benchmark analyses in general and in the specific context of a case study.[3] Not all of the tests or analyses discussed here need to be performed in every damages analysis; the ones that are important for a particular case will depend on the specific facts of the case.

A. Legal Requirements

The legal requirements for methods used to estimate damages, including all econometric methods, fall under the rules for testimony by experts. Under Federal Rule of Evidence 702:

> If scientific, technical, or other specialized knowledge will assist the trier of fact to understand the evidence or to determine a fact in issue, a witness qualified as an expert by knowledge, skill, experience, training, or education, may testify thereto in the form of an opinion or otherwise, if (1) the testimony is based upon sufficient facts or data, (2) the testimony is the product of reliable principles and methods, and (3) the witness has applied the principles and methods reliably to the facts of the case.[4]

Regression analyses have met these requirements many times in litigation for a wide range of issues, including the estimation of antitrust damages,[5] as well as showing at the class certification stage that such

---

3. See Chapter 4, Section C for an overview of the different approaches used to quantify damages, including before-during-after and benchmark (or yardstick) analyses.
4. Fed. R. Evid. 702.
5. *See, e.g.*, Conwood Co. v. U.S. Tobacco Co., 290 F.3d 768 (6th Cir. 2002); *In re* Urethane Antitrust Litigation, 768 F.3d 1245 (10th Cir. 2014).

damages are consistent with the plaintiff's theory of liability.[6] A competent expert should be able to tell the litigant whether sufficient facts and data are available for econometric analysis, explain the types of econometric analyses that are applicable, and reliably apply these econometric techniques.

Using regression analysis does not, by itself, guarantee that an analysis will be viewed as reliable.[7] In general, econometric results will be more reliable as the amount and quality of data increase.[8] If sufficient data are available, econometric analysis has been found necessary to achieve the minimum scientific standard for establishing lost sales and price changes. For example, in *Zenith Electronics Corp. v. WH-TV Broadcasting Corp.*,[9] expert opinion and internal forecasts for sales growth were excluded because data to estimate sales growth via regression analysis were available and not used. The regression must be in a form that assists in determining a material fact, such as the amount of lost sales or the size of price changes.[10] The analysis also must be based on data "reasonably relied upon by experts in the field."[11] Because experts typically verify data for accuracy in consulting work or academic research, experts presenting

---

6. Comcast Corp. v. Behrend, 133 S. Ct. 1426 (2013); *See also In re* Rail Freight Surcharge Antitrust Litigation, 725 F.3d 244 (D.C. Cir. 2013) (remanding in light of *Behrend* so that the district court could analyze whether the regression model proffered by plaintiffs' expert passed sufficient muster at the class certification stage).
7. *In re* Hydrogen Peroxide Antitrust Litigation, 552 F.3d 305 (3d Cir. 2008); Kottaras v. Whole Food Market, Inc., 281 F.R.D. 16, 25-26 (D.D.C. 2012); *In re* Live Action Antitrust Litigation, 863 F. Supp. 2d 966 (C.D. Cal. 2012).
8. Roy J. Epstein, *An Econometrics Primer for Lawyers*, ANTITRUST, Summer 2011, at 29.
9. 395 F.3d 416 (7th Cir. 2005); *see also In re* High-Tech Employee Antitrust Litigation, 289 F.R.D. 555, 582 (N.D. Cal. 2013) (rejecting as evidence proposed factors analysis and compensation charts, but accepting that a regression analysis provided plausible support for the class-wide theory espoused by plaintiffs).
10. Daniel Rubinfeld, *Reference Guide on Multiple Regression*, *in* FEDERAL JUDICIAL CENTER, REFERENCE MANUAL ON SCIENTIFIC EVIDENCE 303 (3d ed. 2011), *available at* http://www.fjc.gov/public/pdf.nsf/lookup/sciman03.pdf/$file/sciman03.pdf.
11. Fed. R. Evid. 703.

regression analyses in court need to conduct similar verifications of the data that they use.[12]

The principle for reliability encompasses many factors entering the regression analysis, and a competent expert should conduct reliability checks to ensure that the results survive the rigors of litigation.[13] Part C of this chapter discusses the basic regression method, and Part D describes many of the design issues that arise in reliably implementing econometric techniques and econometric tests that a testifying expert should perform when appropriate. Moreover, econometric results typically should not change materially with minor changes to the data (e.g., deleting a few observations).[14] Note, though, that while the court decides as a matter of law whether a regression analysis is admissible under the Federal Rules of Evidence, the finder of fact is ultimately responsible for deciding the probative value of such evidence,[15] considering questions including whether the dataset used in analysis is reliable and complete,[16] and, in certain instances, whether statistical significance provides sufficient reliability for the factfinder to base its factual conclusions on the study.[17]

---

12. *See, e.g.*, Maddox v. Claytor, 764 F.2d 1539, 1552 (11th Cir.1985) ("[Multiple regression analysis] measures the probability that the calculated disparity could occur randomly—but the analysis in no way validates the calculation of the disparity itself. If the tested disparity is based on erroneous assumptions or suffers from flaws in the underlying data, then standard deviation analysis is foredoomed to yield an equally faulty result.").
13. Reed v. Advocate Health Care, 268 F.R.D. 573, 593 (N.D. Ill. 2009) ("The issue is not whether [the expert] has shown just any method for proving impact and damages on a class-wide basis; it is whether the method he proposes is a *reliable* means of common proof.") (emphasis in original).
14. Rubinfeld, *supra* note 10, at 199.
15. Cook v. Rockwell Int'l Corp. 580 F. Supp. 2d 1071, 1113 (D. Colo. 2006).
16. *See, e.g.*, Bazemore v. Friday, 478 U.S. 385, 400 (1986) ("Normally, failure to include variables will affect the analysis' probativeness, not its admissibility."); *but see In re* Graphics Processing Units Antitrust Litigation, 253 F.R.D. 478 (N.D. Cal. 2008) (finding missing variables and experts' decision to average certain factors insufficient to carry plaintiff's burden at class certification stage with respect to most of proposed class).
17. Kadas v. MCI Systemhouse Corp., 255 F.3d 359, 362 (7th Cir. 2001) (finding that there is no bright line test for an acceptable significance level, although such a factor still goes towards weight of the evidence). *See also In re* High-Tech Employee Antitrust Litigation, 289 F.R.D. 555, 581 (N.D. Cal. 2013) ("Assuming *arguendo*, that…the [ ] Regression's results are not statistically significant at the 95 percent level does not persuade the Court that the regression is inadmissible (although this failure might affect the

B. Causation and Quantification

Courts typically require an analysis of damages that not only quantifies the amount of damages, but also demonstrates that those damages are causally linked to the allegedly anticompetitive acts.[18] In a price-fixing case, for example, there must be an analysis that provides evidence of a clear link between the agreement to fix prices and an increase in prices that is not explained by other factors (i.e., prices would not have been as high in the counterfactual world in which the agreement had not occurred).[19]

The econometric challenge in establishing and measuring causal effects comes from the fact that most economic data are observational data, derived from observations of market outcomes, not controlled experiments. Therefore, the economist typically has no control over how the data were generated. For example, in a price-fixing case, even if there is variation over time in the existence or extent of collusion, associated variation in prices may or may not capture the *effect* of the cartel, as opposed to the effect of other factors that are changing at the same time. What makes this challenge particularly vexing is that the other changes may *cause* the changes in the existence or extent of collusion. If, for example, collusion is easier to maintain when economic conditions are good, then there will tend to be cartels when prices are high, creating a correlation between high prices and cartel existence even if cartels have little or no actual causal effect on prices.

Another way to state the problem is that identifying causal effects (or "treatment effects" as they are often called in the statistical and econometric literature) requires estimating an unobserved and

---

model's probative value)."); *but see* Allen v. Dairy Farmers of America, Inc., 279 F.R.D. 257, 271 (D. Vt. 2011) (finding plaintiffs' damage model showing a statistical significance with a 99 percent confidence level insufficient to show damage results on classwide basis and denying class certification, because the model lacked a necessary price component).

18. *See, e.g.*, John E. Lopatka, *Antitrust Injury and Causation*, *in* III ABA SECTION OF ANTITRUST LAW, ISSUES IN COMPETITION LAW & POLICY 2299 (2008). *See also* Comcast Corp. v. Behrend, 133 S. Ct. 1426 (2013) (finding that a court must make, as part of its "rigorous analysis," a determination at the class certification stage that plaintiffs' damages theory is consistent with their theory of liability, even if this analysis overlaps with the merits of the claims).

19. Theon van Dijk & Frank Verboven, *Quantification of Damages*, *in* III ABA SECTION OF ANTITRUST LAW, ISSUES IN COMPETITION LAW & POLICY 2335 (2008).

unobservable counterfactual outcome. That is, determining the causal effect of an anticompetitive act requires estimating the difference between the outcomes that were actually observed (e.g., the actual prices during a cartel period) and the counterfactual, "but-for" outcomes that were not observed (e.g., prices that would have been observed absent the cartel), holding all other economic conditions fixed. But if other economic conditions were changing at the same time as the cartel—particularly if those other economic conditions were affecting the likelihood of cartel existence and success—then observed prices in non-cartel periods (often called "clean periods") *do not* provide a good measure of prices during the cartel period but-for the cartel, because in the "clean period" other economic conditions are *not* held fixed.

In statistics, the "gold standard" for answering these kinds of causal questions is the randomized controlled trial (RCT).[20] An investigator conducting a controlled laboratory experiment can set the values of the factor of interest and apply different levels to different test subjects. For example, a sample of subjects in a pharmaceutical clinical trial can be assigned randomly to a "test" group and a "control" group, with the factor of interest (e.g., a pharmaceutical treatment) given only to the test group. Given this experimental design, if, after treatment (or conduct at issue), the outcome of interest (e.g., disease prevalence) changes in the test group in a way that is not found in the control group, then the treatment very likely caused that change and the impact of the treatment can be estimated. Effectively, the experimental design assures that all other conditions are held equal.[21]

The key problem in econometrics, then, is how to identify and measure causal relationships without the ability to use randomization in laboratory

---

20. *See* JOSHUA D. ANGRIST & JÖRN-STEFFEN PISCHKE, MOSTLY HARMLESS ECONOMETRICS 9-18 (2009).

21. In practice, this is generally accomplished by randomization. The role of randomization in this experimental context is not to guarantee that the treatment group and the control group are "matched" on all factors other than the specific treatment being studied; rather, randomization seeks to ensure that there are no systematic differences between the two groups that may affect the outcome of interest other than the treatment. Randomization achieves a kind of matching that recognizes that exact matching is unlikely to be possible in complex situations such as those arising in economics—there are simply too many things that can impact economic outcomes, and it would be difficult, even in experimental settings, to match between treatment and control group on all these factors. *See* RONALD A. FISHER, THE DESIGN OF EXPERIMENTS (1935); RONALD A. FISHER, STATISTICAL METHODS FOR RESEARCH WORKERS (1925).

experiments, but rather based on observational data from real world marketplaces. This problem was discussed at length in a famous paper by Sir Austin Bradford Hill.[22] In this paper, Hill identifies nine criteria for identifying causal effects in observational studies in epidemiology. These nine criteria, while specific to the kind of biomedical data and models arising in epidemiology, also apply in an economic context. Broadly speaking, the Hill criteria mandate that to identify causal effects with observational data, the analysis needs to examine a broad range of implications of the proposed causation: does it make sense in terms of timing, effect size, specificity, mechanism, consistency with economic theory, etc.?

The key point of the Hill criteria is that to reach reliable causal inferences when using observational data, one cannot rely on purely statistical analyses without careful consideration of broader economic ideas. Usually, this approach requires the econometric models to be firmly grounded in relevant economic theory. This basic requirement is also emphasized in the recent work on estimating causal effects by James Heckman.[23]

Keeping these cautions about the difficulties of establishing causal relationships and the limitations of observational data in mind, it is certainly true that, if grounded in economic theory and carefully applied, econometric techniques can provide both reliable estimates of the magnitude of damages and useful information about causation. As discussed above, assuming liability has been established and a reasonable explanatory variable representing the anticompetitive act at issue has been constructed, econometric tests can often reliably be used to determine whether the data are consistent or inconsistent with the anticompetitive act having caused changes in the dependent variable.[24]

While, in principle, it might be possible to match outcomes in circumstances involving the conduct at issue (e.g. cartel behavior) with outcomes in circumstances free of the conduct at issue (e.g. non-cartel

22. A.B. Hill is best known for his work establishing the causal effects of smoking on health using observational data in human populations. The Hill criteria are detailed in A.B. Hill, *The Environment and Disease: Association or Causation?*, 58 P. R. SOC. MED. 295-300 (1965).

23. *See* James J. Heckman, *The Scientific Model of Causality*, 35 SOCIOL. METHODOL. 1-97 (2005). *See also* James J. Heckman, *Econometric Causality*, 76 INT. STAT. REV. 1-27 (2008) (providing additional elucidation and extensions of the role of economic modeling in understanding economic causality).

24. WOOLDRIDGE, *supra* note 1, at 3-4.

behavior) and use such comparisons to draw conclusions about the causal effects of the conduct, such an approach is generally not practical. This matching would need to compare outcomes under the same demand conditions, technology conditions, market structure conditions, and general economic environment between the cartel and the non-cartel circumstances. The possibility of such matching is typically remote because of the number of characteristics involved.

Statisticians have recently proposed an alternative approach to matching for estimation of causal effects: matching on a function (called the "propensity score") that measures the likelihood (or propensity) of receiving the treatment conditional on explanatory factors, rather than matching on all the explanatory factors themselves.[25] Once this match has been made, one can compare outcomes for treated and untreated observations that have similar predicted probabilities of receiving the treatment to infer the causal effects of the treatment.

In some cases, ,there may be difficulties with propensity score matching, however. For example, there may be little overlap between treated and untreated observations in estimated propensity scores: the observations corresponding to the cartel data and the observations corresponding to the non-cartel data often may not seem to match at all in terms of propensity score. In this case, this approach cannot be applied (and, indeed, estimating the causal effect of the cartel activity may be difficult using any statistical methodology).

Because laboratory experiments are generally not an option and the use of propensity scores can be limited, regression models are the more common approach to measuring causal effects of the conduct at issue. As explained above, the regression model allows one to adjust for differences in the economic conditions between the cartel observations and the non-cartel observations by "controlling for," or taking into account the impact of, these differing conditions.

However, these regression models also need to be applied with great care and close linkage to the underlying economic theory. For example, Cochran and Rubin have shown that regression controls can fail to accurately adjust for differing circumstances, if the regression model is not

---

25. Paul R. Rosenbaum & Donald B. Rubin, *The Central Role of the Propensity Score in Observational Studies for Causal Effects*, 70 BIOMETRIKA 41-55 (1983). *See also* Guido W. Imbens, *Nonparametric Estimation of Average Treatment Effects Under Exogeneity: A Review*, 86 REV. ECON. STAT. 4-29 (2004) (providing a more recent review of the literature with a focus on economic applications).

an accurate reflection of the true underlying economic model.[26] A particularly challenging case is when the relationship between the explanatory variables and the outcome of interest varies between the cartel and non-cartel period. Although some adjustments for such differences may be possible, the regression approach fundamentally assumes that the same basic model applies to the data for both the cartel and non-cartel observations—that is, that the economic model is stable across both periods. This may be problematic in markets that are rapidly evolving or when economic conditions are highly unstable over the study period.

In summary, econometric techniques applied to observational data can provide evidence consistent with causality and quantification of impact. The combination of: (1) a properly specified econometric model showing that an explanatory variable has a statistically significant partial effect on the dependent variable, holding constant other factors; and (2) a sound economic theory explaining why one would expect the explanatory variable to have a causal effect, can together provide evidence consistent with the existence of a causal relationship and an estimate of the magnitude of the effect.[27]

For example, suppose a properly specified econometric analysis finds a measure of the effect of the allegedly anticompetitive act on the dependent variable of interest to be statistically significant (discussed below) and of the expected sign.[28] This result would be consistent with the act resulting in damages, as long as there is a clear economic explanation of the linkage between the act and the measure of the partial effect. In this

---

26. *See* William G. Cochran & Donald B. Rubin, *Controlling Bias in Observational Studies: A Review*, 35 SANKHYĀ SER. A 417-46 (1973).

27. *See* Rubinfeld, *supra* note 10, at 184-85. An econometric test called the Granger Causality Test is designed to determine whether changes in one variable occur before changes in another variable. While this property can potentially supply some useful information (for example, one of the Hill criteria is that causal effects of a treatment should not occur before the treatment), it is not the same as causality in the sense in which this chapter uses the term. For example, as PETER KENNEDY, A GUIDE TO ECONOMETRICS 64 (6th ed. 2008), notes, Christmas cards "Granger-cause" Christmas, but obviously do not "cause" Christmas. *See also* Jerry A. Hausman, *Specification and Estimation of Simultaneous Equation Models*, *in* 1 HANDBOOK OF ECONOMETRICS 391, 435-436 (Z. Griliches & M.D. Intriligator eds., 1983) (providing additional discussion of the role of Granger causality in simultaneous equations econometric models).

28. For example, the "expected sign" would be positive if the allegedly competitive act was a price-fixing conspiracy and the dependent variable of interest was price.

case, the econometric estimate can also provide a measure of the magnitude of the impact to estimate the amount of damages. In contrast, a finding that the estimated impact is not statistically different from zero would not provide any statistical support for causation or positive damages, but this lack of statistical support for causation and damages could arise from either lack of actual economic causation and damages or from lack of sufficient information in the observed data.

C. Classic Regression Analysis and Hypothesis Testing

Regression analysis is a statistical method used to analyze the relationship between a dependent variable and a set of explanatory variables based on a sample of observed data. In the context of assessing antitrust damages, the dependent variable is typically an economic outcome important in estimating any damages sustained by the plaintiff. Examples of such outcomes include price, profit margin, and sales quantity. The explanatory variables are typically the major economic factors that may explain variation in the dependent variable.[29] Examples of such economic factors can include consumer income, other consumer demographics that can affect willingness to pay for the product, and manufacturing costs.

Econometric analyses begin with a consideration of the underlying economic forces that generated the data being used to estimate the model. This knowledge should guide the construction of the econometric model, which should incorporate the underlying economic forces.[30] In building the econometric model, the economist makes initial modeling or specification choices by employing economic theory and reasoning. This modeling process identifies the likely major influences on the dependent variable and the potential interactions between those influences. Within the model, the magnitude of the impact of a given economic factor on the dependent variable is typically summarized by a parameter or coefficient that can be estimated from the available data.[31]

These modeling choices are very important. Deviations of the econometric model specification from the true, underlying economic process generating the data can lead to specification error, which make the resulting parameter estimates unreliable—even if they are statistically significant.[32] Fortunately, the validity of many specification choices can

---

29. *See* KENNEDY, *supra* note 27, at 75.
30. *Id.*
31. *See* KENNEDY, *supra* note 27, at 3-4.
32. *See* WOOLDRIDGE, *supra* note 1, at 54-55.

be tested statistically to ensure that a model is correctly specified. This chapter discusses the rationale for some specification choices and how the validity of specification choices can be tested statistically.

1. Basic Structure of a Regression Model

To understand the basics of regression analysis, consider first the relationship between one dependent and one independent variable, known as a simple regression analysis. Assume that prices are determined by cost, plus some fixed mark-up for profits. Most economic models hypothesize a direct positive relationship between cost and price.[33] For purposes of exposition assume a simple model in which price is a linear function of cost though, in practice, more explanatory variables (e.g., those measuring demand) and a more complicated relationship between cost and price can be considered. The deterministic or mathematical relationship between cost and price can be described by the simple algebraic equation of Price $= \beta_0 + \beta_1$Cost.

Here, the *regression coefficient* β_1 indicates how much price will increase with cost.[34] For example, if price (the dependent variable) increases by $0.90 for every dollar that cost (the explanatory variable) increases, then $\beta_1 = 0.9$. A second regression coefficient β_0 is a *constant term* reflecting a fixed markup over changes in costs, say $5.[35] Price will equal the sum of these two effects in this equation. That is if *COST* = $10, then *PRICE* = $5 + (0.9) x ($10) = $14. If *COST* = $12, then *PRICE* = $5 + (0.9) x ($12) = $15.80. This relationship is illustrated in Figure 1.

---

33. Jonathan B. Baker & Timothy F. Bresnahan, *Economic Evidence in Antitrust*, *in* HANDBOOK OF ANTITRUST ECONOMICS 1, 16 (Paolo Buccirossi ed., 2008).
34. G.S. Maddala & Kajal Lahiri, Introduction to Econometrics 59-65 (4th ed. 2009).
35. The "constant term" is sometimes also referred to as the "intercept" because it gives the value of *PRICE* when *COST* = $0 (i.e., the value at which the line Figure 1 crosses or intercepts the vertical axis). Because values of cost (or the explanatory variables included in any regression) are often far from zero, the value of the constant term itself may have little economic significance.

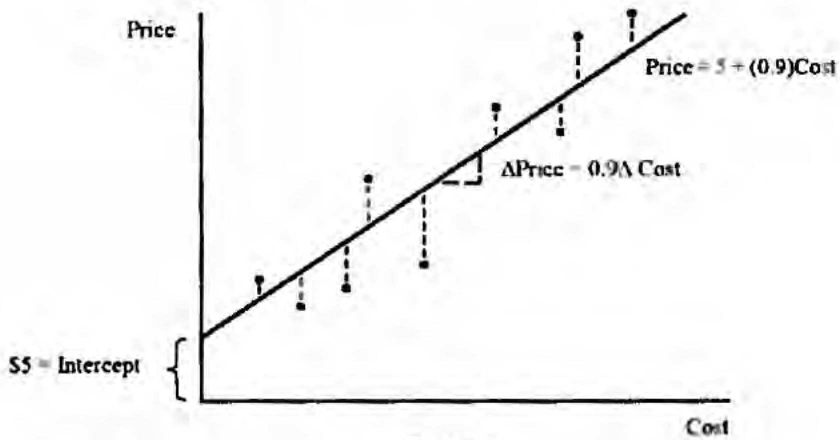


Figure 1

In econometrics, economic relationships are generally considered to be stochastic rather than deterministic. A stochastic relationship is one that has a component that is not generated from a describable systematic process.[36] A stochastic relationship does not yield a single unique value of the dependent variable (e.g., *PRICE*) for a given value of the explanatory variable (e.g., *COST*). In other words, the dependent variable is influenced, but not uniquely determined, by the explanatory variable.

To reflect the stochastic relationship between price and cost, the model can be expanded to be written as Price = $\beta_0 + \beta_1$Cost + ε. The variable ε is known as the "error term."[37] The error term is the difference between the observed dependent variable and the portion of that value that is explained by the explanatory variable.[38] Because the error term is not observed by the econometrician, it amounts to statistical "noise" that partially obscures the underlying relationship between the dependent variable and the

36. KENNEDY, *supra* note 27, at 3.
37. *Id.*
38. By design, the expected value of the error term, conditional on the explanatory variables including the constant term, will typically be zero. In other words, the price of some observations will exceed the predicted price and the price of other observations will fall below the predicted price, but on average the predicted price will match the actual price. *See* GREENE, *supra* note 2, at 20.

included explanatory variable.[39] Specifically, the statistical noise prevents us from being able to measure the regression coefficients β_0 and β_1 from the data exactly.

However, regression analysis provides a way to cut through the statistical noise and obtain *estimates* of each of the regression coefficients, although these estimates will not exactly equal the true underlying coefficients (they are, after all, estimates). Intuitively, one may think of the regression estimate of β_1 as being the average observed impact of cost on price based on the data analyzed, recognizing it is not an exact measure.[40] Under certain conditions the coefficient estimates will have the desirable property of being *consistent*, which means they converge to the true parameter values as the size of the data sample grows.[41]

2. *Multiple Regression*

In reality, there will be multiple influences on the dependent variable or outcome of interest. For example, the price customers paid for a product (PRICE) may be related to three explanatory variables rather than one: (1) the cost of production (COST); (2) the level of industrial production in the downstream industry (one potential proxy for DEMAND); and (3) whether the customers purchased the product during the period of an alleged conspiracy (PERIOD).[42] DEMAND is included because, in

---

39. The error terms may be caused by unpredictable randomness in the behavior of consumers or businesses, the effect of omitted explanatory variables, or measurement error in the dependent variable. *Id.*

40. The coefficient is often assumed to be the same for each unit of observation (e.g., customer or firm). However, an economist may want to test this assumption if there is reason to believe it does not hold. See part D.7. of this chapter for further discussion.

41. *See* WOOLDRIDGE, *supra* note 1, at 56-58. Coefficient estimates from a regression will be consistent if the error term and the explanatory variables are uncorrelated. *Id.*

42. The relevant dependent variable will be determined by the nature of the antitrust allegations, the theory of causation, and specific facts of the case. For example, consider a monopolization case alleging anticompetitive exclusionary conduct through the monopolist knowingly asserting invalid patents against its existing competitors. If before the resolution of the patent cases customers reacted to the pending litigation by switching their purchases to the monopolist (or failing to switch away from the monopolist), the damages to the competitors could in part be the profits on the sales they lost as a result. Given sufficiently reliable data, an econometric analysis of the amount of switching may be relevant both to

addition to cost, economic models in which firms have some degree of market power predict that demand will affect price. PERIOD is included to account for the alleged anticompetitive act. For instance, the allegation in most price-fixing conspiracies is that average prices are higher during the conspiracy period, all else equal.[43]

The relationship between the dependent variable and the explanatory variables might take the following form: Price = β_0 + β_1Cost + β_2Demand + β_3Period.

The coefficient associated with each explanatory variable indicates the amount by which the dependent variable changes when the associated explanatory variable changes, holding the other explanatory variables constant.[44] Thus, the regression coefficients (β_0, β_1, β_2, and β_3) are interpreted as the partial effect of each explanatory variable on PRICE. In this example, β_0, as before, is the constant term; β_1 is the coefficient associated with COST; β_2 is the coefficient associated with DEMAND; and β_3 is the coefficient associated with PERIOD.[45]

Again, suppose β_1, the coefficient associated with COST, is equal to 0.90. In a multiple regression context this implies that the customer's price would be higher by $0.90—holding DEMAND and PERIOD constant—if costs increased by $1.[46] Further assume that β_2 equals 0.10, so that increasing DEMAND by 1 unit would increase price by $0.10, holding COST and PERIOD constant. PERIOD takes on the value of one in the alleged conspiracy period and zero at all other times (i.e., before or after the alleged conspiracy period). Similar to the other coefficients, β_3 measures the impact of PERIOD on price, holding COST and DEMAND constant. This example assumes that β_3 is 0.50, so that this impact is $0.50 for the alleged conspiracy period, meaning that, holding all other factor equal, prices were $0.50 higher during the conspiracy period than before or after the conspiracy period.

---

causation issues and to the quantification of damages. Also note that a regression model used in practice often will have more than three explanatory variables. A relatively simple model is used here for expository purposes.

43. See Chapter 8 for an overview of the legal and economic issues involved in estimating damages associated with anticompetitive price increases or overcharges. As discussed below, the use of a single "dummy" variable for the conspiracy period may not be appropriate. *See* part E of this chapter.

44. WOOLDRIDGE, *supra* note 1, at 14-16.

45. MADDALA & LAHIRI, *supra* note 34, at 127-132.

46. This result is sometimes described by saying that the pass-through rate equals 0.9.

To reflect the stochastic relationship between price and the three explanatory variables, the model can be written as: $\text{Price} = \beta_0 + \beta_1\text{Cost} + \beta_2\text{Demand} + \beta_3\text{Period} + \varepsilon$. The error term ε contains pure underlying statistical noise. In addition, the error term also contains the influence of other factors that may be at work but are omitted from the model. As a practical matter, the set of explanatory variables included in an econometric regression model never accounts for all of the factors that affect the dependent variable.[47] The variables that are not included in the model are collectively subsumed in the model's error term.

In this example, the unobserved factors that appear in the error term might include particular aspects of the supply and demand conditions facing the customer, such as the end uses to which the customer puts the product, the availability of substitutes for the product (which may differ across customers), the production level of the particular industry in which the customer participates (since the included demand variable DEMAND may not be industry-specific), the customer's negotiating skill if prices are individually negotiated, and many other such factors specific to the circumstances at hand.[48]

3. *Estimation Methodology*

The damages expert will have to select an appropriate methodology for estimating the model parameters and the associated standard errors (a measure of how precise the estimated parameters are, which is used in testing hypotheses about the parameters and determining the margin of error around the parameter estimates, as described in more below). The most frequently used method for estimating the coefficients that summarize the relationship between the explanatory variables and the dependent variable is *ordinary least squares* (OLS).[49] This statistical technique minimizes the squared deviations between data points and a line passing through those data points, thus creating the line that fits the pattern of the data as closely as possible, based on this squared-deviation standard

47. KENNEDY, *supra* note 27, at 3.

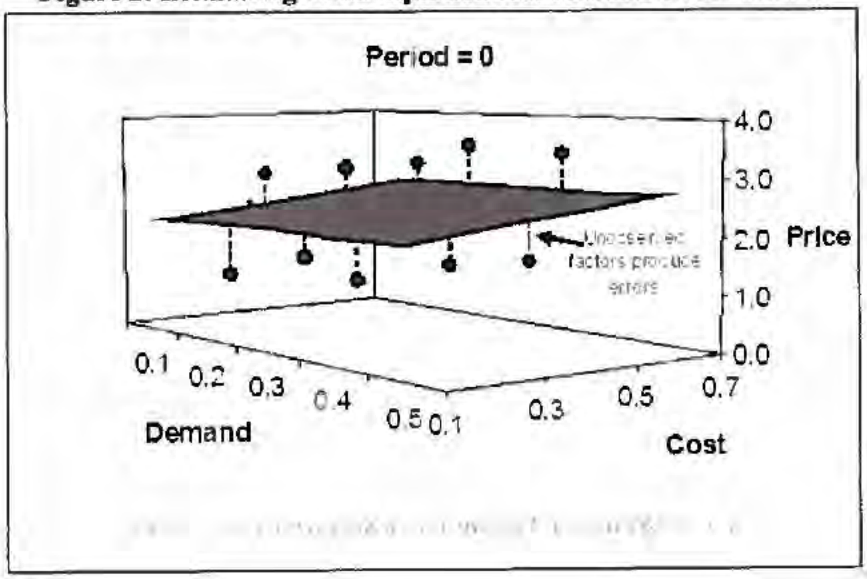
48. *See, e.g.*, John H. Johnson & Gregory K. Leonard, *Economics and the Rigorous Analysis of Class Certification in Antitrust Cases*, 3 J. COMP. L. & ECON. 341, 346-47 (2007).

49. Other estimation techniques use different criteria to determine the coefficient estimates. Examples of other techniques include generalized method of moments and maximum likelihood estimation. *See* WOOLDRIDGE, *supra* note 1, at chs. 12-14.

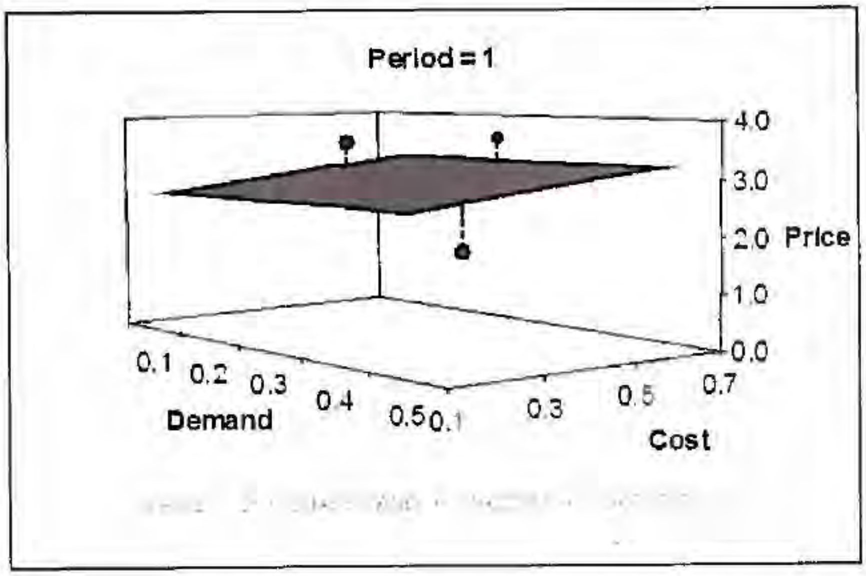
for closeness of fit.[50] This technique has the desirable property of being consistent under general conditions.[51]

Figure 1 (above) shows the regression line in relation to the actual observations of prices and costs used in the regression. Similarly, Figure 2 (below) shows the planes discussed above within the actual observations used in the regression. The coefficients of the equation are estimated by minimizing the squared errors between the data points and the plane in order to get the "best fit" of the plane to the available data.

Figure 2: Estimating Price Equation from Actual Observations



50. KENNEDY, *supra* note 27, at 48.
51. WOOLDRIDGE, *supra* note 1, at 56-58. OLS moreover will be *unbiased*—correct on average in small samples—under the condition that the conditional expectation of the error term given the explanatory variables is zero. *See* GREENE, *supra* note 2, at 63-65. Because the concepts of consistency and unbiasedness have similar meanings for purposes of this chapter, the terms "consistent" and "unbiased" are used interchangeably.



Estimating the standard errors of the parameter estimates requires that the properties of the regression's error term be properly taken into account. For example, standard errors frequently are estimated assuming the error term of the regression is *independently and identically distributed* across the observations used to estimate the regression model. That is, the error for each observation is assumed to be from the same distribution (or pattern) of possible errors, and each error from each observation is statistically independent of the others (the value of the error for one observation contains no information about the distribution of the error for another observation).[52] If, instead, the error terms are not independent (e.g., correlated with each other) or not identically distributed, then the resulting standard error estimates generally will be inconsistent unless this more complicated error distribution is accounted for.[53] We discuss approaches to estimating standard errors given more complex error structures in Part D below.

4. *Evaluating the Reasonableness of Estimation Results*

Before using the estimation results to opine on the impact of the alleged anticompetitive conduct, experts typically subject their estimation results to some basic sanity checks. The appropriate checks will vary from

52. *Id.* at 54-57.
53. *Id.*

case to case, depending on the circumstances. However, two types of features experts typically consider are: (1) the reasonableness of the signs of the estimated coefficients; and (2) the ability of the model to predict or "fit" the data.

Regarding the first—reasonableness of the signs—one might ask whether the cost and demand factors have the predicted signs. All else equal, economic theory predicts that as costs increase or as demand for a product increases, price should increase. Thus, theory generally predicts that the estimated coefficients on the cost and demand factors should have positive signs (i.e., take on positive values). However, with many variables in a regression model, it is not uncommon for particular coefficients to have unexpected signs; hence, while it is valuable to check whether the overall pattern of coefficient signs makes economic sense, econometric models need not be rejected simply because a small number of coefficients (out of a larger set) have unexpected signs.

Regarding the second—the ability of the model to fit the data—most regression software packages generate so-called goodness-of-fit measures. These measures describe how well a statistical model fits a set of data. This discussion concentrates on the most common goodness-of-fit measure, *R-squared* (or R^2). R^2 is the percentage of the total variance in the dependent variable that is explained by the regression model. A higher R^2 means that the model explains more of the variance in the dependent variable. In the extreme, if the model were a perfect fit, the R^2 would attain its maximum possible value of 1.[54] If the model explained none of the variation in the dependent variable, the R^2 would attain its minimum possible value of 0.

A very small R^2 may raise concern about whether the model is missing important explanatory variables.[55] However, one must not ascribe too much importance to R^2, because a high R^2 (close to 1) is not required for a regression model to be reliable. Indeed, if one is interested in estimating the partial effects of the explanatory variables on the dependent variable,

---

54. KENNEDY, *supra* note 27, at 14. R-squared can be increased by simply adding variables to a regression, even if they are not statistically significant. To account for this problem, the *adjusted R-squared* is sometimes calculated. Adjusted R-squared increases when a new variable is added to the regression only if that new variable has a t-statistic greater than one. *Id.* at 94.

55. For a discussion, see Rubinfeld, *supra* note 10, at 179, 188, 216-17. However, even with a small R^2, least squares produces unbiased estimates, as long as the conditional expectation of the error term given the included explanatory variables is zero.

R^2 is generally irrelevant as long as the omission of other exogenous variables does not create any bias in the estimation of the coefficient for the variable of interest (which would occur if, for example, the omitted variables are uncorrelated with the variables included in the regression model).

In deciding whether one regression model is "better" than another at predicting the outcome of interest, one might want to compare R^2 across two regression models. However, one must also take care in so doing. Such a comparison can only be given any weight if the two regression models have the same dependent variable. Otherwise, the comparison is one of apples to oranges. For example, consider one model where the dependent variable is average list price and a second model where the dependent variable is the average transactional price (including discounts). Even if the first model explains a larger percentage of the variation in price, the first model is not necessarily "better" than the second model if the regression is intended to estimate the impact of an act on the prices customers pay. Moreover, even when the dependent variable is the same in two regressions, the model selection criteria discussed below are likely a better basis for comparison than R^2.

5. Statistical Inference: Hypothesis Testing and Confidence Intervals

Upon estimating the regression model, the expert will use the estimated results for statistical inference. Statistical inference refers to hypothesis testing designed to probe the validity of the allegations regarding the conduct at issue. Statistical inference relies upon the standard errors and confidence intervals associated with the regression estimates. As such, we turn here to a discussion of these topics.

The starting point for hypothesis testing is that there are economic statements or arguments that the finders of fact will want to evaluate. These are conclusions such as "the cartel had no appreciable effect on prices" or "this antitrust violation elevated prices in this market at this time." As discussed above, economists use models that simplify and focus the complex economic interactions that take place in markets to analyze such economic questions. Typically, the complex economic questions that are of interest to the finders of fact are reduced to statements about the values of parameters in these models. Thus, for example, the economic question of interest, "the cartel had no appreciable effect on prices," is represented by the statistical hypothesis that the coefficient on the PERIOD variable is zero.

To evaluate this question statistically, the hypothesis test will consider whether the estimate of the PERIOD coefficient obtained from the data is consistent with what would be expected if in fact the true PERIOD coefficient were zero. As discussed above, the estimated coefficient for PERIOD likely would not be exactly equal to zero even if the true underlying coefficient were zero. However, the estimated coefficient should not be "far" from zero (measured relative to the parameter's standard error) if the hypothesis were true.

The usual framework used for hypothesis testing is called the Neyman-Pearson framework.[56] This framework is built on the following simple structure. Any statistical hypothesis can either be true or false, and the finder of fact can either reject the hypothesis as inconsistent with the data or accept the hypothesis (or more formally, fail to reject the hypothesis) as consistent with the data. This framework leads to two possible kinds of errors that the finder of fact can make: she can accept a hypothesis that is false, or reject a hypothesis that is true. Ideally, of course, the econometrician would like to minimize the chance of either kind of error, but the essential tension is that any stringent criterion for accepting a hypothesis (i.e., a criterion that demands a high level of evidence to accept the hypothesis) will have a higher chance of rejecting the hypothesis when it is true. Similarly, a stringent criterion for rejecting will lead to a higher chance of accepting when the hypothesis is false. The Neyman-Pearson approach, which is the most widely used approach in statistics and statistics in legal settings, is to limit the chance of rejecting the hypothesis when it is true to some small number, typically 5 percent (this is called the level of the test), and then, conditional on this level, to design the statistical procedure to minimize the chance of accepting the hypothesis when it is false (that is, to maximize the power of the test).

Ideally, the economist would choose the level of the test to balance the costs of falsely rejecting a true hypothesis with the costs of failing to reject a false hypothesis. However, such an analysis is rarely carried out. Instead, statistical hypotheses are carried out at conventional levels of 5 percent, occasionally at 1 percent or 10 percent. These choices for the level of the test should be understood as conventions and not necessarily grounded in any careful balancing between the kinds of errors that can occur in practice.

This approach, as typically implemented, rejects hypotheses only if there is *substantial* evidence against the hypothesis. This approach is explained by the fact that these methods were developed to assist scientists

56. GREENE, *supra* note 2, at 111.

in analyzing experimental data. Scientists may not want to reject hypotheses until there is overwhelming evidence that they are inconsistent with experimental evidence. However, care must be taken in applying these methods in other situations, such as legal proceedings. The question must be asked: Does the asymmetric nature of the hypothesis test reflect the question that the finder of fact needs to answer? Does this asymmetry in any way tilt the analysis of the data?

Implementing a hypothesis test typically requires a measure of the *statistical precision* of the estimate of the coefficient. If there is a great deal of unexplained variation in the data (referred to as statistical noise), the coefficient estimate will be highly imprecise and provide relatively little information in testing a hypothesis about the true value of the coefficient.[57] As an example, an opinion poll based on a very small sample of respondents generally would not be very precise and thus would not provide a very useful basis for testing a hypothesis concerning the percentage of the overall population who held an opinion. A larger sample of respondents would produce a more precise estimate, but getting information from the entire population is often not necessary to get a reasonably accurate estimate.

a. Confidence Intervals

To measure the precision of a statistical estimate, econometricians and statisticians typically use what is called a standard error.[58] The standard error may be best understood by explaining how it is used to create a *95 percent confidence interval*.[59] For many statistical estimates, a 95 percent confidence interval is approximately equal to the range defined by the coefficient estimate plus and minus two times the standard error. One can be "95 percent confident" that the true underlying coefficient will lie within the appropriately constructed 95 percent confidence interval. Economists often use 95 percent confidence intervals, but (much like they sometimes use different levels for hypothesis tests) they also sometimes

57. KENNEDY, *supra* note 27, at 67-68.
58. WOOLDRIDGE, *supra* note 1, at 44.
59. Kaye & David A. Freeman, *Reference Guide on Statistics*, *in* FEDERAL JUDICIAL CENTER, REFERENCE MANUAL ON SCIENTIFIC EVIDENCE 211 (3d ed. 2011), *available at* http://www.fjc.gov/public/pdf.nsf/lookup/SciMan3D01.pdf/$file/SciMan3D01.pdf.

use intervals defined by lower or higher levels of confidence, such as 90 percent or 99 percent.[60]

When the coefficient estimate is relatively imprecise, the 95 percent confidence interval will be relatively wide, reflecting the greater uncertainty regarding the true value of the coefficient. Conversely, when the coefficient estimate is relatively precise, the 95 percent confidence interval will be narrower.[61] Returning to the opinion poll example, pollsters often report the 95 percent confidence interval, e.g., "45 percent of respondents supported the proposal with a margin of error of +/- 5 percent." That is, intervals constructed this way will contain the percentage of the population supporting the proposal about 95 percent of the time.

The margin of error (i.e., the width of the confidence interval) shrinks, all else equal, as the sample size of the poll increases and the precision of the poll increases accordingly. With a margin of error of +/-1 percent, the 95 percent confidence interval for the percentage of the population supporting the proposal would be narrower, from 44 percent to 46 percent.

b. T-statistics

Standard errors also can be used to conduct hypothesis tests regarding coefficients of interest. Suppose the economist wishes to test the hypothesis that the coefficient on the PERIOD variable is equal to zero. To test this hypothesis, the economist could calculate the ratio of the coefficient estimate to its standard error, or "$\hat{\beta}/\text{se}(\hat{\beta})$" where $\hat{\beta}$ is the coefficient estimate and se($\hat{\beta}$) is the standard error of the coefficient estimate. This ratio is called a *t-statistic*.[62]

If the hypothesis is correct and the true underlying coefficient is in fact zero, then the t-statistic should not be very far from zero. If the t-statistic turns out to be far from zero, it would cast doubt on the truth of the hypothesis. How do we determine whether the t-statistic is "far" from zero? We can calculate the probability that the t-statistic achieves a certain value, if the hypothesis were true. For example, if the hypothesis were true, there is about a 90 percent probability that the t-statistic will fall between 1.7 and -1.7 and about a 95 percent probability that the t-statistic will fall between 2 and -2. Thus, if the hypothesis were true, there would be only a 5 percent probability that the t-statistic we observe would be either greater than 2 or less than -2. Accordingly, if we observe a t-statistic greater than

60. GREENE, *supra* note 2, at 1061.
61. KENNEDY, *supra* note 27, at 54.
62. GREENE, *supra* note 2, at 115-17.

2 or less than -2, the data would appear to be inconsistent with the hypothesis (because such an outcome is quite unlikely if the hypothesis were in fact true).

Indeed, if the absolute value of the t-statistic that the economist calculates exceeds two, then the hypothesis that the true underlying coefficient equals zero typically would be said to be *rejected* at the 5 percent significance level and the result typically would be termed *statistically significant*.[63] This result often is also expressed by saying that the coefficient is "statistically significantly different from zero (at the 5 percent level of significance)." As noted above, the 5 percent level of significance (and the corresponding 95 percent confidence interval) is often used by economists and statisticians when conducting hypothesis tests, but other levels of significance, such as 1 percent or 10 percent, are also sometimes used.

As an example of these techniques of statistical inference, suppose that the coefficient estimate on the PERIOD variable in the price regression was 0.50, which would imply that prices were $0.50 higher during the alleged conspiracy period as compared to outside that period, holding constant the variables COST and DEMAND (and assuming correct model specification). Suppose further that the standard error of the coefficient estimate on PERIOD is 0.20. In this case, the 95 percent confidence interval would be approximately $0.10 to $0.90.

Similarly, the standard error can be used to test the hypothesis that the coefficient on the PERIOD value is zero, which implies that price was not higher during the alleged conspiracy period (holding constant COST and DEMAND). This test would be conducted by calculating the t-statistic: 0.50/0.20 = 2.50. Since the calculated t-statistic of 2.50 is greater than two, the coefficient on PERIOD would be said to be statistically significantly different from zero at the 5 percent level of significance, and the hypothesis that the price was no higher during the alleged conspiracy period would be rejected.

c. P-values

The results of statistical tests often appear to be a simple yes-no answer, but this may be an inadequate summary of the statistical information available in hypothesis tests and hence make it difficult for the judges and juries to appropriately weigh this evidence with the other kinds of evidence. A more fine-grained summary of the information that measures the strength of the evidence can be helpful.

63. *Id.*

A more fine-grained summary of the information in a hypothesis test comes from the *p-value* for the hypothesis. The p-value represents the level of significance that the hypothesis can be rejected at. Thus, a p-value of 0.05 would indicate that the hypothesis could be rejected at the 5 percent level we have been discussing. In our example, if the coefficient on the PERIOD variable were 0.5 and the standard error were 0.2, as above, the t-statistic for testing the hypothesis would be 0.2/0.5 = 2.5. The details of calculating a P-value are a bit more complicated than calculating a t-statistic and cannot be performed efficiently within this text. However, a p-value will be reported by any statistical software used to run the regression and will be easily available. In this case, the p-value corresponding to the 2.5 t-statistic would be 0.012, which means that the hypothesis could be rejected using even more stringent criteria than the usual 5 percent level for testing (the 1.2% level). On the other hand, suppose that the coefficient were 0.5 and that the standard error were 0.4, then the t-statistic would be 0.5/0.4 = 1.25. In this case, the p-value would be 0.21; that is, the hypothesis would only be rejected under the less stringent criterion that the chance of wrongly rejecting a true hypothesis could be as large as 21 percent. While an economist may refer to the simple yes/no dichotomy of whether the test passes at the 5 percent level, the p-value can provide more detail about how close this test was. In general, small values of the p-value indicate that the test provides a substantial amount of evidence that is inconsistent with the null hypothesis and large values of the p-value indicate the opposite. However, it is important not to interpret the p-value as the chance that the hypothesis is true; the hypothesis itself is either true or false; in contrast, p-values are chances of observing stronger (in repeated random samples) evidence against the hypothesis, assuming that it were true.

The example of the last paragraph illustrates another important point in interpreting hypothesis tests. In both examples, the values of the estimated coefficients on the PERIOD variable have the same value: 0.5. In one case, the hypothesis that the cartel had no appreciable impact on prices is strongly rejected (p-value of 0.012), and in the other example, the hypothesis cannot be rejected at conventional levels of significance (p-value = 0.21). However, the actual best estimate of the price impact of the cartel is the same in both cases. In the model as specified, the estimated coefficient would represent an economically significant effect of the cartel as the best estimate that prices were higher by $0.50 in both cases, even though in one case one could not say with assurance that the effect was different from zero. This stresses a critical point about hypothesis tests: they represent statistical statements about both the estimated coefficient

and the precision with which it is estimated, not purely economic statements about the size of the measured effect.

It is also possible to test hypotheses concerning more than one estimated parameter.[64] For example, suppose that, rather than a single PERIOD variable to capture the cartel period, the model had included PERIOD1 and PERIOD2 variables that identified subperiods that might have had different overcharges. The economist could test the hypothesis that the coefficients on PERIOD1 and PERIOD2 are both zero, that is, that the cartel had no impact on prices even allowing for different effects in different subperiods. This kind of test is usually carried out by using an F-statistic, which measures how far the two estimated coefficients are from zero, again measured relative to the precision with which these coefficients are estimated. The form of this test is more complex than the t-test, but the basic testing logic is the same.

D. Necessary Decisions and Common Areas of Expert Debate

In practice, experts will need to make a series of design decisions to build an appropriate econometric model and a series of implementation decisions to arrive at reliable estimates to answer the questions of interest. Many of these decisions will inevitably become the subject of expert debate. This section discusses some of these decisions.

1. Choice of Explanatory Variables

As discussed above, the explanatory variables in an econometric model represent economic factors that influence the dependent variable.[65] An important question is which explanatory variables to include in the model. Answering this question should begin with economic theory combined with knowledge of the industry. For example, if the dependent variable is price, economic theory suggests that cost factors are likely the most important explanatory variables, with demand drivers and industry capacity as potentially important explanatory variables as well, among other things.[66] Industry knowledge would suggest specific variables that would appropriately represent these factors. If a product is used as an input

---

64. JAMES H. STOCK & MARK W. WATSON, INTRODUCTION TO ECONOMETRICS 219-223 (3d ed. 2010).
65. *See* part A of this chapter.
66. Jonathan B. Baker & Daniel L. Rubinfeld, *Empirical Methods in Antitrust Litigation: Review and Critique*, 1 AMER. L. & ECON. REV. 386(1999), at 391.

by downstream industries, the level of production in those industries might drive the demand for the product.[67] For costs, there may be direct measures of production costs from firms producing in the industry (perhaps the defendants in a price-fixing case), and the prices of the inputs used to produce the product could also be useful measures of costs.

a. Too Many or Too Few Variables?

The number of explanatory variables suggested by economic theory and industry knowledge often will be large. Is it best to include all of the explanatory variables, or should one try to pare back the number of variables in order to have a simpler model? The downside to including extraneous explanatory variables in a regression is that the coefficients may be less precisely estimated. However, these estimates will still be unbiased.[68] The effect on precision of having additional variables will often be small when the sample size is large, but the potential loss of precision for smaller samples counsels careful consideration of which explanatory variables to use in those circumstances.

Mistakenly excluding important explanatory variables in an attempt at simplicity, on the other hand, can result in an *omitted variable bias* and misinterpretation of estimated results. As an example, suppose one omits from a regression model a control for consumer income. Suppose further that during the period of conspiracy, consumer income happened to increase. In this case, the regression model will predict that prices were higher in the conspiracy period. Yet, one might falsely conclude from these results that prices were higher in the conspiracy period *because of*

---

67. Care needs to be taken that the explanatory variables used in a least squares regression are *exogenous*, or uncorrelated with the error term, to the extent possible. *See* WOOLDRIDGE, *supra* note 1, at 54-55. For example, if the intermediate product in question represents a large share of the downstream industry's costs, the amount of downstream production may be affected by the price of the intermediate good. If the impact of the price of the intermediate good on the sales of the downstream product is substantial, then downstream production could be considered *endogenous* (correlated with the error term) rather than exogenous. *See* Baker & Rubinfeld, *supra* note 66, at n.17. Methods of detecting and dealing with endogeneity are discussed later in this chapter.

68. More specifically, ordinary least squares is still the best linear unbiased estimator as demonstrated by the Gauss-Markov Theorem, one of the more famous theorems in statistics. *See* GREENE, *supra* note 2, at 60. This means that not only is least squares unbiased, but also it is the most efficient (i.e., most precise) among linear unbiased estimators.

the conspiracy. However, some or all of the estimated conspiracy effect may be attributable to the omitted income variable.

More precisely, omitted variable bias arises when important explanatory variables that have been omitted from the regression model are correlated with included explanatory variables. Because the omitted variables are in the error term, the result will be a correlation between the included explanatory variables and the error term.[69] In the language of econometrics, this correlation with the error term is a problem known as endogeneity.[70] This misspecification will bias the resulting coefficient estimates, and make these estimates unreliable for damage estimation. This bias does not diminish as sample size increases.[71]

To see why, suppose that in the regression model discussed above, there is a second important demand driver variable, DEMAND2, that was omitted from the regression model. Suppose further that this variable was positively correlated with the alleged conspiracy period variable PERIOD. For example, DEMAND2 might be unusually high during the period of the alleged conspiracy, so DEMAND2 and PERIOD would be highly positively correlated. All else equal, when DEMAND2 was unusually high, PRICE should be unusually high. A regression model that included PERIOD, but did not include DEMAND2 would mistakenly attribute the effects of DEMAND2 (which was omitted from the model) to PERIOD, the variable that represents the alleged anticompetitive act. As a consequence, the regression would lead to the mistaken conclusion that the alleged conspiracy represented by PERIOD was the "cause" of the high prices, when in fact the cause was the unusually high values of DEMAND2 during the alleged conspiracy period. This bias will only occur, however, if DEMAND2 is empirically important for the dependent variable PRICE and is also correlated with PERIOD.

In general, one should be more concerned with avoiding bias than with improving precision. That is, it is better to be on average right though imprecise, than to be precisely wrong. For these reasons, it is in general better to include more variables than fewer.[72] Nonetheless, no variable should be included unless there is a solid economic rationale for putting it in the model. And, in general, one should test model results to ensure that they are not overly sensitive to the inclusion of a particular explanatory variable, particularly if the economic rationale for including that variable is questionable.

---

69. WOOLDRIDGE, *supra* note 1, at 65-67.
70. We address methods to identify and deal with endogeneity below.
71. *Id.*
72. KENNEDY, *supra* note 27, at 95.

When there is concern that extraneous variables could greatly affect the precision with which coefficients can be estimated, a parsimonious model may have virtue.[73] There are several model selection criteria one can use to compare models containing different sets of explanatory variables.[74] Examples are the *Akaike Information Criterion* (AIC) and the *Schwartz Criterion* (SC).[75] These criteria reward increased explanatory power of the model, and penalize the addition of extraneous variables. When the penalty from the addition of a sufficiently weak explanatory variable exceeds the reward, the criteria will indicate that the variable should not be added. There are other sophisticated econometric methods that attempt to explicitly trade off bias against increased precision of the estimates.[76] However, these techniques should still only consider variables that economic theory and industry information identify as likely to be important, and in no case should these techniques be used in a mechanical way or in a way that substitutes for sound economic reasoning.

b. Multicollinearity

Multicollinearity occurs when two or more of the explanatory variables are highly correlated with each other ("collinear").[77] When this happens, the coefficient estimates on the collinear variables may be imprecise; that is, they may have large standard errors. It then becomes difficult to make precise statistical inferences about the true underlying coefficients. The coefficient estimates may also have the wrong sign or be sensitive to small changes in model specification.

Dropping one or more of the collinear variables is sometimes argued to solve the multicollinearity "problem." This proposition is incorrect and

---

73. *Id.*
74. *Id.* at 206.
75. *Id.*
76. Such techniques are often called "shrinkage" estimators. *See* KENNEDY, *supra* note 27, at 198. Because these techniques lead to biased estimators, there is no consensus among economists as to whether they should be used. *Id.* at 201.
77. Correlation coefficients measure how closely two data series are related to each other. When the two series form a perfectly straight line in a positive relationship ($y = \alpha x$), the correlation coefficient will be 1. When the two series form a perfectly straight line in a negative relationship ($y = -\alpha x$), the correlation coefficient will be -1. More generally, when the two series do not form a perfectly straight line but have random variations from a straight line ($y = (+/-)\alpha x + \varepsilon$), the correlation coefficient will be between -1 and 1. *See* GREENE, *supra* note 2, at 1031-32.

can lead to a significant bias in the estimates of the coefficients on the retained variables.[78] Put differently, dropping one of the collinear explanatory variables does not solve the core problem that the effects of the two collinear variables on the dependent variable are hard to disentangle and thus imprecisely estimated. Rather, dropping one variable simply arbitrarily sets its coefficient to zero and assigns all of the effect to the other variable.

Dealing with multicollinearity requires recognizing several important points. First, multicollinearity is a matter of degree. Explanatory variables are almost always correlated with each other to some extent. Since multicollinearity is a characteristic of the sample of data rather than a characteristic of the underlying population, there is no rigorous statistical test for multicollinearity.[79] A number of indicators of multicollinearity have been developed, but many of these have serious shortcomings.[80]

Second, multicollinearity among a subset of the explanatory variables may affect the precision with which the coefficients on those explanatory variables are estimated, but not affect the precision with which the coefficients on other explanatory variables are estimated.[81] Accordingly, even severe multicollinearity among a subset of explanatory variables—those that are included as "controls" but not as the main variables of interest—may not affect the coefficient estimates that are of most interest in the analysis. In a price-fixing case, for example, interest typically focuses on the coefficient on the alleged conspiracy variable. Multicollinearity that affects the precision with which the demand variable coefficients can be estimated should be of little concern if it does not affect the precision with which the alleged conspiracy variable coefficient can be estimated. For this reason, a useful indicator for multicollinearity is that proposed by Belsley, Kuh, and Welsch.[82] This indicator focuses on

78. KENNEDY, *supra* note 27, at 197.
79. DAVID A. BELSLEY *ET AL.*, REGRESSION DIAGNOSTICS: IDENTIFYING INFLUENTIAL DATA AND SOURCES OF COLLINEARITY 95 (1980). Statistical tests allow one to make an inference about a population characteristic from the sample; but multicollinearity is a sample characteristic, not a population characteristic. While there may be some true correlation in the population, this true correlation is not the issue with multicollinearity. What matters with collinearity is the correlation in the particular sample that one has to work with, so testing what is "true" in the population is uninteresting.
80. KENNEDY, *supra* note 27, at 195.
81. BELSLEY ET AL., *supra* note 79, at 91.
82. *See id.* Another indicator is the "variance inflation factor." *See* GREENE, *supra* note 2, at 90.

identifying which explanatory variable coefficient estimates are potentially affected by the multicollinearity.

Third, if the precision of the coefficient estimate on the variable of primary interest (e.g., the alleged conspiracy variable) is not substantially affected by the multicollinearity, there is no reason to drop any variables that have a strong economic rationale for inclusion. As discussed above, dropping variables that are truly important could lead to omitted variable bias. A large change in the alleged conspiracy variable coefficient estimate after dropping collinear variables could indicate omitted variable bias, which would render the estimate of the alleged conspiracy variable coefficient unreliable. Therefore, in general the best course of action in that case would be to retain the collinear variables.[83]

Now assume the variable of interest is affected by the multicollinearity. Would it now be valid to drop the other presumably economically important variable(s) with which the variable of interest is collinear? Generally no valid basis exists for choosing to retain the variable of interest as opposed to retaining the variable(s) with which it is collinear. As explained above, the multicollinearity and the imprecision of estimation that results correctly indicates that it is difficult to separate out the individual effects of the collinear variables.[84] In that situation, there may be no "solution" other than gathering more data, which often may not be feasible. With severe collinearity between the variable of interest and other economically important explanatory variables, the coefficient on the variable of interest cannot be estimated precisely, and the regression model may therefore be unreliable for estimating damages. Dropping explanatory variables to reduce the multicollinearity without a sound economic reason amounts to assuming away rather than fixing the problem.[85]

2. *Functional Form*

The relationship in the simple examples relating prices to costs and other factors assumes that the explanatory variables have a linear (i.e., straight line) relationship with the dependent variable. In particular, a given change in the dollar value of cost translates to a fixed dollar change in price. However, economic theory, statistical evidence, or information

---

83. KENNEDY, *supra* note 27, at 197.
84. Jacques Dreze, Comment, *Nonspecialist Teaching of Econometrics: A Personal Comment and Personalistic Lament*, 2 ECONOMETRIC REVIEWS 291, 296 (1983), states this view quite well when he says that omitting a variable in this situation "amounts to elevating ignorance to arrogance."
85. *Id.*

from market behavior may indicate that a different functional form for the regression equation may be more appropriate. For example, in certain instances a given percentage increase in costs will yield a proportionate percent increase in price, rather than a fixed dollar relationship. In this case, a regression equation similar to the one described above can be estimated, but the price and cost data would be transformed into logarithms and the cost coefficient would (approximately) measure the impact of a percentage change in costs on the percentage change in price.[86] If possible, there should be an economic basis for the functional form of the regression equation that is estimated, although economics often provides only a rough guide for the appropriate functional form, in which case it may be appropriate to try alternative functional forms to ensure that the ultimate answer to the economic question of interest is relatively stable across alternatives (that is, that the results are "robust" to different functional forms).[87]

3. *Structural Models versus Non-Structural Models*

Econometric models can generally be divided into two broad types: *structural models* and *non-structural models*.[88] A structural model consists of equations reflecting the relationships between the variables in the model that are directly derived from economic theory.[89] For example, the structural model for price and quantity in an industry might consist of a demand equation—representing how customer demand for the product is determined—and a supply equation—representing how firm supply of the product is determined. Price and quantity would be the endogenous variables, since they are determined by the other variables in the model. By contrast, a non-structural model consists of equations reflecting the relationships between outcomes (like price and quantity) and only exogenous supply and demand factors that influence the outcomes, through an economic model.

86. KENNEDY, *supra* note 27, at 106.

87. *Id.* at 102.

88. *See* KENNEDY, *supra* note 27, at 171-72. *See also* Peter C. Reiss & Frank A. Wolak, *Structural Econometric Modeling: Rationales and Examples From Industrial Organization*, *in* HANDBOOK OF ECONOMETRICS, VOL 6A, CHAPTER 64, 4277-4415 (J.J. Heckman and E. Leamer eds., 2005) (providing additional discussion of the distinction between structural and non-structural models).

89. *Id.*

Structural models have certain advantages over non-structural models. In particular, because the structural model is derived from economic theory, the exercise of implementing the structural model requires the analyst to be explicit about the assumptions undergirding the model. Making the underlying assumptions explicit facilitates a discussion about whether those assumptions are reasonable given the question being addressed and the facts of the industry being studied. Non-structural models also require such assumptions but they are often implicit rather than explicit, which can make assessing and debating those assumptions more complicated.

The use of structural modeling confers other potential advantages. One such advantage is that the imposition of an economic structure on the econometric specification allows the analyst to directly estimate economic parameters of interest. For example, the coefficient on price in the demand equation would be related to the elasticity of demand (the percentage change in demand for a given percentage change in price holding all other drives of demand fixed). Another advantage of structural modeling is that it facilitates analysis of counterfactuals—a key consideration in the estimation of antitrust damages. For example, a structural model could be used to investigate the impact of the entry of a new competitor on economic outcomes such as price or output.

Nonetheless, structural equations can be difficult to estimate. Thus, experts often estimate a non-structural or "reduced form" model, instead of a structural one.[90] The reduced form is an algebraic rearrangement of structural models so that each equation has an endogenous variable on the left side and only exogenous explanatory variables on the right side. For example, the analysts might specify a demand model and a supply model. Each model is a function of price and quantity, both of which are endogenously determined. In addition, it can be assumed that the quantity demanded will equal the quantity supplied (though this need not be the case if, for example, regulations such as price limits interfere with the workings of the market). This setup generates three equations (the demand equation, the supply equation, and the identity equation equating quantity demanded to quantity supplied) and three unknowns (equilibrium price, quantity demanded, and quantity supplied).

Furthermore, depending on the application, the reduced form model may be more directly useful even if estimating the underlying structural model is feasible. This is because the reduced form model can provide the net effect of the conduct at issue on the outcome of interest (like price).

---

90. Baker & Rubinfeld, *supra* note 66, at 391-92.

By net effect, we mean that the estimated reduced form coefficients take into account both supply and demand side responses. By contrast, with a structural model, one would have to calculate the net effect given the estimated parameters.

To see this, we observe that one can start with an explicitly specified structural model and then solve it to obtain the reduced form model, or one can directly specify the reduced form model. In this latter case, the endogenous variables price and quantity are expressed in terms of the exogenous variables and the error terms, and the coefficients on the exogenous variables in the reduced form are functions of (depend on) the structural parameters.[91] Thus, the coefficients of reduced form models are generally a mixture of the structural parameters from the underlying structural equations (if they are derived from structural equations at all) and thus often do not have a direct economic interpretation.[92]

Should one use a structural model or a non-structural model? If the question of interest centers on the value of a structural parameter (e.g., the elasticity of demand), estimation of a structural model is generally the more appropriate way to proceed.[93] However, if the question of interest concerns determining the expected value of an outcome variable given the values of the exogenous variables, estimation of a non-structural model is generally a more straightforward way to proceed, provided sufficient data are available.[94]

For damages analysis in antitrust cases, the question of interest often involves estimating the likely "but-for" or counterfactual value of an outcome variable, such as price.[95] Accordingly, non-structural models will often be better suited to antitrust damages analyses. Even if one is going to estimate a non-structural model, however, consideration of the basic elements of the appropriate structural model can be very useful in identifying the appropriate explanatory variables to include in the reduced form model since the reduced form model is obtained by solving the structural model for the endogenous variables in terms of the exogenous variables.[96]

---

91. *See* KENNEDY, *supra* note 27, at 171-72.
92. *See* Baker & Rubinfeld, *supra* note 66, at 392.
93. *Id.* at 405.
94. *Id.* One can predict the endogenous variables from a structural model by solving the model given the estimated structural parameters. However, this method is generally not as straightforward as estimating the reduced form model directly. *Id.* at 414-16.
95. *Id.* at 389.
96. *Id.* at 391.

4. Estimation of Parameters Using Instrumental Variables

The estimation of structural relationships—for example, a relationship between the outcome of interest and another outcome that itself was jointly determined within an economic model—requires more information than the estimation of reduced form relationships. In such cases, the ordinary least squares or OLS method described above can yield biased and inconsistent estimates.[97] As a matter of principle, the coefficient estimates from an OLS regression will be consistent only if the error term and the explanatory variables are uncorrelated. This condition will not be satisfied if there are "endogenous" explanatory variables or, in other words, explanatory variables that are themselves jointly determined with the outcome variable.

Possibly the most famous example of such a situation arises in the context of studying the relationship between price and quantity. Because price and quantity are jointly determined within an economic model, the regression of price on quantity will not be appropriate unless one faces special circumstances. Otherwise, the simple OLS approach will not be appropriate.

More specifically, suppose one had access to observations of pairs of price and quantity across markets. Each price and quantity pair represents the equilibrium outcome in the market. The analyst would like to use the variation in price and quantity to trace out the demand relationship between price and quantity. However, without more information, this exercise would be impossible because price and quantity are jointly determined by demand *and* supply conditions. Put simply, the analyst would not know if she were tracing out a demand curve, a supply curve, or (the most likely case) some combination of the two.

For another example, consider a regression of price on market concentration as measured by the Herfindahl-Hirschman Index (HHI)—a standard measure of market concentration—designed to test the proposition that greater concentration leads to higher prices. While this regression could be estimated using OLS techniques, such an approach would not be appropriate if the concentration measure was correlated with the error term. Such correlation could arise, for example, if more concentrated industries have higher costs and sufficient data to control for those costs are unavailable.[98] Because costs are likely to impact prices in

---

97. WOOLDRIDGE, *supra* note 1, at 89.

98. Timothy F. Bresnahan, *Empirical Studies of Industries with Market Power*, *in* HANDBOOK OF INDUSTRIAL ORGANIZATION, VOL II, CHAPTER 17, 1011-1057 (R. Schmalensee and R.D. Willig eds., 1989).

unobserved ways and those costs are also correlated with the degree of concentration in the industry, the concentration measures will be endogenous.

To estimate the parameters of a structural model, one needs to introduce additional information. This additional information, referred to as an "instrumental variable" or "IV," must satisfy two conditions: (1) it must be uncorrelated with the error term, meaning it does not belong in the regression model; and (2) it must be correlated with the endogenous explanatory variable (after netting out the effects of the other explanatory variables).[99]

Returning to the price-concentration example above, a commonly used instrument in that context is the number of firms competing in the market.[100] Depending on the specifics of the market(s) in the study, the number of firms in the market is likely to be correlated with other measures of concentration, but uncorrelated with the error term in the pricing equation. Anything that shocks prices—such as demand or cost shocks—will also affect shares and thus concentration ratios. But as long as those shocks do not drive firms into or out of the market then they do not determine the number of firms (e.g., under certain economic conditions, the number of firms in the market may be less likely to be correlated with measures of cost). If these conditions are met, then the number of firms competing in the market is a valid instrument and an IV regression would yield consistent estimates of the coefficients.

Similarly, in the case of estimating a demand curve, IVs that shift the supply curve (e.g., exogenous cost shocks across markets) could serve as valid instruments because they will be correlated with the equilibrium price (through the supply conditions) and uncorrelated with quantity demanded. Thus, IVs that shift the supply curve help the analyst trace the demand curve.[101]

There are several econometric approaches to IV estimation. One widely used implementation technique is known as "two-stage least squares" (or 2SLS to distinguish it from OLS regressions).[102] An

99. WOOLDRIDGE, *supra* note 1, at 89-121.

100. Craig Peters, Evaluating the Performance of Merger Simulation: Evidence from the U.S. Airline Industry, XLIX J. LAW ECON. 627-649 (2006).

101. Analogously, if the analyst was interested in tracing out the supply curve, IVs that shift the demand curve would be important.

102. In the first stage, the analyst regresses the endogenous explanatory variable on all exogenous explanatory variables, including the instrumental variables. In the terminology of the structural models discussed in the prior section, this first-stage regression can be thought of as a reduced form

alternative and more general IV estimator is called the generalized method of moments (GMM).[103] Common to these methods is the requirement of a sufficient number of valid instrumental variables or IVs.

Relatively recent advances in the estimation of demand models help to further illustrate the themes discussed here. In many industries, market-level price and quantity data are available along with the characteristics of products from which consumers can choose. Product characteristics may include price along with non-price quality characteristics. For example, an automobile may be described as a combination of a price, manufacturer (e.g., Ford, Toyota, etc.), vehicle type (e.g., sedan, light truck, etc.), fuel efficiency, and a term reflecting unobserved quality of the automobile. By making certain structural assumptions related to the functional form of consumer utility, the analyst can relate consumer choices to the aggregate product-market-level data in order to estimate the parameters of the demand model. Starting from this basic theoretical model of consumer demand, the economics literature has derived reduced form econometric specifications that relate functions of aggregate market share to linear combinations of product characteristics.[104]

The model just described introduces concerns about the endogeneity of price (and perhaps certain non-price product attributes as well). For example, the econometric error term, which represents at least in part unobserved product quality, is likely to be correlated with both prices and market shares. Products with relatively higher unobserved quality are likely to have both higher market share and higher prices, all else equal. Failing to address this issue is likely to lead to underestimates of the price elasticity of demand (i.e., the responsiveness of demand to changes in price).

Several types of IVs may be appropriate in this scenario. First, as discussed above, cost shifters, if available, typically provide valid instruments. A second set of instruments follows from an economic model of differentiated product competition. In such a setting, the product

---

equation. In the second stage, the fitted or predicted value of the endogenous explanatory variable is used as an explanatory variable in place of the endogenous explanatory variable itself. The second-stage model can be estimated using OLS regression techniques. *See* WOOLDRIDGE, *supra* note 1, at 89-121.

103. WOOLDRIDGE, *supra* note 1, at 207-238.

104. See, *e.g.*, Steven T. Berry, *Estimating Discrete-Choice Models of Product Differentiation*, 25 RAND J. ECON. 242-262 (1994); Steven Berry, James Levinsohn, and Ariel Pakes, *Automobile Prices in Market Equilibrium*, 63 ECONOMETRICA 841-890 (1995) [hereinafter *BLP*].

characteristics of rival products (e.g., the number and quality of rival products) may offer valid instruments.[105] This conclusion follows from the identifying assumption that rival product characteristics are uncorrelated with the unobserved product quality.[106] A third set of instruments consists of the prices of the products in other markets, which may be assumed to be correlated with prices in the market of interest through common cost structures, but uncorrelated with the market-specific product valuations.[107]

5. *Estimation of the Standard Errors*

As described above, correct statistical inference (hypothesis testing and formation of confidence intervals) requires not only estimates of the coefficients of the model, but also estimates of the standard errors of these coefficient estimates.[108] For example, a t-test of whether a coefficient β is zero is conducted by forming the t-statistic "$\hat{\beta}/\mathrm{se}(\hat{\beta})$," discussed in the last section. This t-statistic can be invalid and lead to incorrect statistical inference if either the coefficient estimate *or* the standard error of the coefficient estimate, $\mathrm{se}(\hat{\beta})$, is inconsistently estimated.[109]

The textbook approach to standard error estimation is to impose the assumption that the errors of the regression are *independently and identically distributed* across the observations used to estimate the regression model. If, instead, the error terms are not independent (e.g., they are correlated with each other) or not identically distributed, then the resulting standard error estimates generally will be inconsistent unless this more complicated error distribution is accounted for.[110]

Correlation among the errors of different observations can arise in various situations. For example, suppose the data sample is a *time series*, i.e., the data were generated by observing the variables (e.g., price, cost,

105. See Timothy Bresnahan, Departures from Marginal-Cost Pricing in the American Automobile Industry, 17 J. OF ECONOMETRICS 201-227 (1981); Timothy Bresnahan, Competition and Collusion in the American Automobile Oligopoly: The 1955 Price War, 35 J. OF INDUSTRIAL ECONOMICS 457-482 (1987); BLP, supra note 106.
106. *Id.*
107. Such instruments would not be valid in the presence of common demand shocks across markets.
108. WOOLDRIDGE, *supra* note 1, at 57.
109. As noted above, "inconsistent" is a technical term that means, roughly, that even if the econometrician had enormous amounts of data—perfect information—the estimate would not coincide with the true value that it is purporting to measure.
110. *Id.*

and demand) at various points over time (e.g., on a monthly basis).[111] In such a case, the error in one month might well be related to the errors in adjacent months, since the unobserved economic factors that appear in the error term might themselves exhibit correlation over time. This correlation of errors over time is called *serial correlation*.[112]

As another example, suppose the data sample is a *cross-section/time series*, or *panel data* set, where the variables are observed at various points of time separately for each of a number of units of observation, such as individual customers.[113] Each customer's data are a time series. Therefore, the error terms for a given customer may exhibit serial correlation. In addition, each customer may have idiosyncratic factors that affect the price it paid, but that are unobserved in the data. These factors would be present in all of the errors across time for that customer, which would be a further cause of correlation among the errors for a given customer. This effect is called an *unobserved individual-specific effect* (where the "individual" refers to the unit of observation, e.g., a customer).[114]

Importantly, the correlation among errors need not be confined to errors that pertain to the same customer. For example, the error terms for all customers within the same time period also may be correlated. Unobserved economic factors may affect all customers' prices at a given point in time and therefore these common factors will appear in the errors of all of the customers in a given time period. Similarly, if these unobserved factors are themselves serially correlated, then the error for one customer in one month will be correlated with the error for another customer in another month. Therefore, there may be correlation among the errors both within and between units of observation in a panel data set.[115]

There can be important consequences from estimating the standard errors for the coefficient estimates as if the errors were uncorrelated when, in fact, they are correlated. In this situation, a statistical test on the coefficients may yield what appears to be a statistically significant result even though it is not.[116] To see why this is so, suppose that the estimate of

111. *See* GREENE, *supra* note 2, at 1048.
112. *Id.* at 903.
113. *Id.* at 1048.
114. *See* WOOLDRIDGE, *supra* note 1, at 282.
115. This problem is widely recognized in the econometrics literature. *See* Brent R. Moulton, *An Illustration of the Pitfall in Estimating the Effects of Aggregate Variables on Micro Units*, 72 REV. ECON. STAT. 334 (1990); Marianne Bertrand et al., *How Much Should We Trust Differences-in-Differences Estimates?*, 119 Q. J. OF ECON. 249 (2004).
116. *Id.*

the coefficient on a price-fixing conspiracy indicator variable (i.e., a variable that takes a value of one during the period of the alleged conspiracy and zero otherwise) is 0.15. Further suppose that the incorrectly estimated standard error, ignoring the correlation among the errors, is 0.05, but that the correctly estimated standard error, taking into account the correlation among the errors, is 0.15. To test the hypothesis that the alleged conspiracy had no effect, the expert should calculate a t-statistic of 0.15/0.15, using the correctly calculated standard error of 0.15. Because the correct t-statistic is less than 2, the hypothesis that the alleged conspiracy had no effect would not be rejected at conventional levels of statistical significance. However, suppose that the expert uses the wrong standard error. In this case, the (incorrect) t-statistic would then be 0.15/0.05 = 3, and the hypothesis that the alleged conspiracy had no effect on prices would be strongly rejected. Thus, use of the incorrectly calculated standard error could lead to the wrong inference.[117]

Econometricians have a variety of methods for consistently estimating the standard errors when correlation among the errors exists. In a time series context (discussed in more detail below), various non-parametric procedures (procedures that do not impose any functional form on the correlation) have become widely used.[118] In a panel data context (also discussed in more detail below) these procedures may also be used, and

---

117. Once again, this stresses the point that hypothesis testing depends on the precision of the estimates, not just the value of the coefficients. All that changed when the serial correlation was accounted for was that the precision of the estimate was revealed to be lower, with no change in the value of the estimated coefficient. Yet this change caused the estimate to become statistically insignificant, meaning that the hypothesis that its true value was zero could not be rejected.

118. See Donald W. Andrews, Autocorrelation and Heteroskedasticity Consistent Covariance Matrix Estimation, 59 ECONOMETRICA 817 (1991); Donald W. Andrews & J. Christopher Monahan, An Improved Autocorrelation and Heteroskedasticity Consistent Covariance Matrix Estimator, 60 ECONOMETRICA 953 (1992). The resulting standard errors are also consistent in the presence of heteroskedasticity (the variance of the error term differs across observations). Many statistical software packages implement the Newey-West procedure, which is one example of such a procedure to obtain autocorrelation and heteroskedasticity consistent standard errors. See Whitney Newey & Kenneth West, A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix, 55 ECONOMETRICA 703 (1987).

other methods have been proposed as well.[119] Some of these panel data procedures are easily implemented.[120]

These procedures produce consistent estimates of the standard errors even when there is no correlation among the error terms. In other words, they work well both when the error terms are independent and identically distributed and when they are not. Thus, these procedures have become used more generally in practice.[121] When there is good reason to suspect the existence of correlation among the errors, such procedures should be used to avoid making incorrect statistical inferences.[122]

Finally, another feature of the error term—known as heteroskedasticity—can also create problems in accurately measuring standard errors. Heteroskedasticity occurs when the variance of the error term varies across observations, even if the error terms are independent.[123] This condition is another violation of the independent and identically distributed error term assumption (in this case, the error distribution is not identical across observations) that can cause the traditional standard error calculation to be inconsistent. Again a well-known and widely used technique exists for calculating standard errors that are robust to heteroskedasticity (White standard errors).[124] This technique is also readily available in many econometric packages.[125]

The standard error estimation methods discussed above all rely on mathematical models and approximations for the statistical procedures that are used to construct the coefficient estimates and for the distributions of the error terms. These mathematical models and approximations typically work well when the statistical procedures are well-behaved and the datasets are large. However, there are many situations where one or both of these conditions may not be met.

---

119. Bertrand et al., *supra* note 115.
120. For example, Stata, a popular econometrics software package, includes a "cluster" option for calculating standard errors assuming unspecified within-group (cluster) correlation between the error terms. *See* STATA PRESS, BASE REFERENCE MANUAL 81 (2007).
121. GREENE, *supra* note 2, at 920, n.10.
122. If one is not sure, there are various tests one can run to test the hypothesis of no correlation in the error terms. *See* WOOLDRIDGE, *supra* note 1, at 130, 279, 420-449; GREENE, *supra* note 2, at 922-924.
123. GREENE, *supra* note 2, at 257.
124. Hal White, A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity, 48 ECONOMETRICA 817, 838 (1980).
125. For example, Stata includes an option for calculating White standard errors. *See* STATA, *supra* note 120, at 81.

An alternative approach to the estimation of standard errors that can work quite generally has been developed recently in the statistics and econometrics literature. This method is called *bootstrapping*.[126] The basic idea behind bootstrapping is quite simple: after carrying out the estimation of the model, the econometrician constructs a large number of "replication samples" of the data by resampling from the original data. Estimating the same model using these replication samples and measuring the variability of the resulting estimates can yield accurate and reliable estimates of the standard errors. Bootstrap methods are a useful and increasingly used tool.

Bootstrapping works well in many, but not all, situations. However, for the bootstrap method to take into account serial correlation, cross-sectional correlation, or heteroskedasticity, the resampling method used to construct the replications samples needs to take these potential sources of error correlation into account. For example, if there is concern that the errors may exhibit within-customer correlation, then the replication samples should be chosen by randomly sampling customers. Or if the data are time series data and there is concern that the errors may be serially correlated, the replication samples should be based on the block-bootstrap, in which the data are resampled as contiguous blocks of observations.

6. *Misspecification Bias and Specification Tests*

Regression analysis provides unbiased estimates of the regression model coefficients generally only if the regression model is correctly specified and if the error term (the unexplained component of the dependent variable) is not correlated with the explanatory variables.[127] If these conditions do not hold, the regression model may not produce reliable coefficient estimates. In such cases, the model will not reliably measure the effect of the allegedly anticompetitive act on the variable of interest. *Thus, from an economic and statistical point of view, it is important to test whether a regression model is misspecified in order to avoid drawing incorrect or unreliable conclusions.*[128]

Misspecification can arise from,[129] among other reasons, the imposition of incorrect restrictions on the regression coefficients, the

---

126. *See, e.g.*, Joel L. Horowitz, *The Bootstrap*, *in* HANDBOOK OF ECONOMETRICS, VOL 5, CHAPTER 52, 3159-3228 (J.J. Heckman and E. Leamer eds., 2001).
127. WOOLDRIDGE, *supra* note 1, at 53-58.
128. GREENE, *supra* note 2, at 479-80.
129. Another source of bias, the omission of important explanatory variables from the model, is discussed in part D.5.a. of this chapter.

assumption of an inappropriate functional form, or inclusion of endogenous variables as explanatory variables.[130] As noted above, endogenous variables are variables that are correlated with the model's error term, in some cases because they are caused, in whole or in part, by the model's dependent variable. For example, the quantity sold of a good is determined in part by its price. Thus, quantity is typically an endogenous variable in a regression model where price is the dependent variable. In other cases, endogeneity can arise from omitted variable bias, since an omitted variable is part of the model's error term and may also be correlated with one or more of the model's explanatory variables.

In many circumstances, these problems can be remedied if detected and if the appropriate data are available.[131] Fortunately, there are well-known and generally accepted econometric tests for the presence of misspecification.[132] These tests often provide an objective and scientific means to determine whether a regression model is reliable.

For example, omitted variable bias can be tested by identifying and including in the regression model additional explanatory variables that economic reasoning and other market information suggest are likely to affect the dependent variable. If these additional explanatory variables turn out to be statistically significant, and the coefficient estimates on the previously included explanatory variables change substantially when the additional variables are added, then the regression model that omitted the additional explanatory variables likely is misspecified and its results are biased and unreliable.[133] More generally, results should be tested to make sure they are robust to reasonable changes in the set of control variables—including reasonable alternative variables used to capture similar economic concepts, such as alternative cost or demand controls. Models that yield results that are highly sensitive to particular choices of explanatory variables have a high likelihood of being affected by specification error and thus should generally not be relied upon.

Another potential misspecification is caused by imposing incorrect restrictions on the coefficients of the regression model—that is, by forcing the coefficients to take incorrect values. Forcing incorrect values on some

130. KENNEDY, *supra* note 27, at 54-55.
131. WOOLDRIDGE, *supra* note 1, at 54, 55, 62-63.
132. *Id.* at 46; *see also* Jerry A. Hausman, *Specification Tests in Econometrics*, 46 ECONOMETRICA 1251 (1978) (providing additional discussion of the use of specification tests).
133. A.H. Studenmund, Using Econometrics: A Practical Guide 168-175 (6th ed. 2011).

coefficients can bias other coefficients.[134] For example, suppose that the data contain information on two different customers (A and B), and that the regression model imposes the restriction that both customers were affected in the same way by the alleged conspiracy.[135] Further assume that customer A has a perfect substitute to which it could turn in response to an anticompetitive price increase, while customer B has no substitute. In that case, even if there were a conspiracy, the alleged conspirators would be able to impose an anticompetitive price increase only on customer B. The existence of a substitute for customer A would prevent the alleged conspirators from imposing a supracompetitive price increase on customer A. If the coefficient on PERIOD were allowed to be different for the two customers, customer A would have a zero coefficient (since its price would not be higher during the alleged conspiracy period), while customer B would have a positive coefficient, reflecting say a $1 overcharge. However, a regression model that forced the PERIOD coefficient to be the same for the two customers likely would result in a positive coefficient estimate that was biased for both customers (too high for customer A and too low for customer B). For example, the estimate of the PERIOD coefficient might reflect only a $0.50 overcharge. Forcing the PERIOD coefficients to be the same for the two customers could lead to the incorrect inference that both customers were harmed by the alleged conspiracy, when in fact customer A was not harmed. This restriction may also significantly affect the accuracy of the estimate of total damages. For example, if customer B had purchased 100 units at $11 and customer A purchased 10 units at $10, the true total damages would be ($1*100 + $0) = $100. Yet, the results from the regression model in this example would lead to a total damages estimate of approximately ($0.50*100 + $0.50*10) = $55, a substantial underestimate. The existence of bias resulting from the imposition of incorrect restrictions can be econometrically tested using standard procedures often used by economists.[136]

Another kind of misspecification that can lead to flawed conclusions is the restriction that the model be the same across different periods of time.[137] For example, if there is a pre-cartel period and a post-cartel period bracketing the cartel period, then it would be desirable to use both time

134. *Id.* at 121-122.

135. This is an example of an assumption of "structural stability"—that the model is "stable" across the groups of customers, meaning that the same model holds for both groups.

136. STUDENMUND *supra* note 133, at 121-127.

137. This is another form of a structural stability assumption, in this case stability over time.

periods in the damages model. However, if there were significant changes in the competitive structure of the market between the pre- and post-cartel period, the pooled estimates may form a misleading benchmark or clean period.

A standard method for detecting changes in model parameters is the Chow test.[138] The Chow test is more accurately described as a testing principle than a single test, and it describes methods to test if all the coefficients in the model have changed between two periods (or subgroups of the observations), or if some specific subset of parameters has changed, or if some particularly important combination of parameters (or function of the parameters) has changed.[139] Of course, the specific form of the Chow test should be specified in advance, so that the econometrician does not go on a data-mining exercise to find any kind of change. Such a data-mining exercise would lead to incorrect statistical inferences using the test.

One special case of "data-mining" that has been formally developed as a valid testing methodology is procedures to identify structural breaks that occur at unknown times. The idea is that the economist can perform Chow tests sequentially over the time period spanned by the data and identify the point in time at which a structural break occurred.[140]

E. Additional Considerations for Specific Forms of Data

Different types of data can also require special consideration. This section discusses issues that tend to arise in the context of large administrative data, panel data, and time series data, respectively.

1. Estimation with Large Datasets

An important source of data for analyzing antitrust damages claims are the administrative records kept by firms. These can include detailed transaction-level data that can amount to millions of individual records. These records typically record information on individual transactions by customer and by specific product. Frequently, there can be hundreds or thousands of customers and hundreds or thousands of specific products,

---

138. GREENE, *supra* note 2, at 168-70.
139. STOCK & WATSON, *supra* note 64, at 226-228.
140. *See, e.g.*, Donald W.K. Andrews, *Tests for Parameter Instability and Structural Change with Unknown Change Point*, 61 ECONOMETRICA 821-856 (1993); Jushan Bai & Pierre Perron, *Estimating and Testing Linear Models with Multiple Structural Changes*, 66 ECONOMETRICA 47-78 (1998).

all with differing characteristics. These administrative records can potentially provide a wealth of data to assess antitrust claims, but the availability of such "big data" does not guarantee that reliable estimates of antitrust damages can necessarily be easily or readily uncovered.

There are a number of considerations in working with large administrative datasets:

1) The data were collected for business purposes, and may not include information on explanatory variables that are relevant for estimating antitrust damages.
2) The data may come from multiple administrative units that use different formats, track different information, or encode information in their data in different ways. For example, the multiple firms involved in a price-fixing cartel may use different systems to track their transactions, the same firm may change systems during the conspiracy period, or different divisions within a firm (e.g., sales department and finance department) may track information differently. In addition to the practical concerns, this can create the potential for "structural breaks" or heterogeneity unrelated to actual economic behavior.
3) Units of observation in administrative data may not align with economically relevant events. For example, invoicing data often records rebates, discounts, and returns separately from the underlying transaction in a way that makes it difficult to align these records. In such a case, it may be difficult to compute accurate prices for these transactions.

For these reasons, the promise of "big data" available in administrative records needs to be viewed with an appropriate amount of caution, and the analysis of these data requires careful attention to both the underlying market economics generating the data and details of how the data are collected, processed, and retained.

Properly used, administrative data may allow the econometrician to study whether the data are consistent with the full gamut of economic implications of the alleged anticompetitive behavior. As emphasized by both Fisher and Hill (as referenced above), assessing whether observational data are fully consistent with theoretical predictions is an important component of identifying causal effects when using such data.

For example, data that is rich in detail on customers and on individual products would allow the estimation of models that measure anticompetitive effects at the customer and product level. The

econometrician can determine if the pattern of estimated effects is consistent with the alleged violations (e.g., some customers may not have been affected by a price fixing conspiracy because the terms of their sales were determined by a contract that was agreed to before the start of the alleged conspiracy). If not, the pattern of estimated effects would suggest that the allegations in the case are incorrect or that there were no anticompetitive effects. Further, if the estimates indicate that some products or customers benefited (e.g., the estimated but-for prices are higher than the actual prices) from the alleged anticompetitive acts, again there would be concern about the validity of the allegations, assuming the model was reliably estimated.

There are, however, a number of statistical concerns when building econometric models with highly detailed and heterogeneous transaction level data including:

1) Models need to account for potential differences in economic outcomes (such as prices) across the customers, products, time periods, and so on. These differences are often modeled by the use of "fixed effects," which are discussed in a subsequent section. However, simple inclusion of such fixed effects may not account for all the heterogeneity in the data. For example, perhaps the underlying variation in the outcome is such that including fixed effects for each customer-product combination is economically or statistically important. If so, the model may require estimation of many more parameters than is possible to estimate reliably even with a large dataset.
2) It may be difficult to estimate accurate standard errors because the patterns of potential correlations in the error terms across observations are difficult to model. While clustered standard errors may allow the econometrician to address some of these issues, if the patterns of correlations in the error terms potentially reach across all the dimensions of the data, then the required clusters may be too large to be statistically valid. In some cases, bootstrapping techniques, discussed above, can help to compute more reliable standard errors.
3) Transaction level data may be irregularly observed over time in ways that make estimation of time series or panel data models computationally intractable.

For all these reasons, it can be advantageous to aggregate administrative data before conducting econometric analyses. For example,

the econometrician may aggregate sales for groups of customers together or aggregate sales for a particular customer in a month together. Of course, any such aggregation should reflect the underlying economics of the case and should be careful not to eliminate econometrically or economically important variation or heterogeneity in the underlying data. There are no bright lines in balancing the benefits of preserving this variation with the potential complications of dealing with the disaggregated data, and so using detailed administrative data effectively requires careful economic and statistical analysis to achieve this balance.

2. *Estimation with Panel Data*

Datasets can be composed of single observations over a period of time (a *time series*) or of a number of observations in single time period (a *cross section*). In general, data sets consisting of both time series and cross-sectional observations are called cross-section time series data—or panel data.[141]

Panel data allow for richer econometric specifications than either time series or cross-sectional data.[142] Consider for example, a dataset that tracks customer transactions. If there were 100 transactions in each year (corresponding to a purchase by each of 100 customers) over five years, the data would allow the econometrician to study both within-customer and across-customer behavior.

There may be reasons to believe that there are unobserved factors specific to each customer that affect the dependent variable (e.g., price) in such a model. For example, suppose some customers are better negotiators than others. Good negotiators are likely to receive better prices on each transaction than poor negotiators. Similarly, suppose some customers have better substitutes for the product in question than other customers. Then, the customers with good substitutes are likely to receive better prices on each transaction than customers with poor substitutes.

When these *unobserved individual-specific effects* are present, the error term of the regression can be represented as composed of two parts.[143] The first part represents the unobserved individual-specific effect, which is assumed to be fixed across all purchases of a given customer. The second part of the error term can change from one time period to the next

141. For a discussion of issues involved in using panel data, see part D.5 of this chapter.

142. Two often-used econometric specifications for panel data are *fixed effects* and *random effects* models. *See* WOOLDRIDGE, *supra* note 1, at 321-34.

143. *Id.* at 285.

for each customer. How one proceeds in this situation depends on whether the unobserved individual-specific effects are correlated with the explanatory variables.

If the unobserved individual-specific effects are correlated with the explanatory variables, ignoring them will create an endogeneity problem, which in turn will bias the regression estimates.[144] Because there are multiple observations for each customer over time in panel data (but not a cross-section), one can control for the individual-specific effects, and thus eliminate the endogeneity concern. The easiest way to do so is to include in the regression specification a set of customer indicator variables—one for each customer.[145] The dummy variable for customer *i* will control for the unobserved individual-specific effect for customer *i*—in effect, it reflects the average value of the dependent variable for customer *i*, controlling for other observable factors, where the average value for each individual may be different. This type of analysis is called a "*fixed effects regression*," where the fixed effect for customer *i* is equal to the regression coefficient on its unique dummy variable.[146]

For an example of how failing to control for fixed effects can bias a study's results, consider a study of the effects of education on individuals' earnings. An individual's motivation and inherent ability are difficult to measure fully with observable variables. In an earnings regression model, motivation and inherent ability therefore will appear in the error term as an individual-specific effect that will increase the individual's earnings over his or her lifetime. Yet these same factors are also likely highly correlated with the individual's education. Without including fixed effects for the individuals in the regression, a regression of earnings on education (and other explanatory variables) may well produce biased results, since the coefficient on education would then also reflect the effects of motivation and inherent ability on earnings.[147] A regression that excludes individual fixed effects will rely on variation across individuals as well as over time for specific individuals to identify the effect of education on earnings, thereby subjecting itself to the concern that education and unobserved individual characteristics may be correlated across individuals. On the other hand, a regression that includes individual fixed

144. *Id.* at 300-15.

145. Here, an indicator variable for customer *i* will equal one for that customer, and zero for all other customers. *See* MADDALA & LAHIRI, *supra* note 34, § 8.2.

146. *See* WOOLDRIDGE, *supra* note 1, at 300-315.

147. *Id.* at 65.

effects will rely only on variation over time for each individual to identify the effect of education on earnings, thereby lessening this concern.

If the unobserved individual-specific effects are not correlated with the explanatory variables, a fixed effects model still yields reliable results, but a better estimator may be available. In such cases, what is known as the "*random effects estimator*" may provide more statistically precise results than fixed effects. However, before relying on random effects results, it is critical to test the validity of the assumption of no correlation between the unobserved individual-specific effect and the explanatory variables.[148]

3. *Estimation with Time Series Data*

In a time series context, model specification and estimation proceeds differently depending on whether the time series in question have *stochastic trends* or not.[149] The term stochastic refers to the unpredictability of a random variable over time. A series with a stochastic trend not only trends upward or downward, but shocks to the series persist for a long time.[150] Economic time series that are not stochastically trending will sometimes be referred to as *stationary,* while econometric time series that are stochastically trending will sometimes be referred to as *non-stationary.*

Stochastic trends can create problems in econometric estimation. For example, suppose a dependent variable is increasing at the same time an explanatory variable is increasing. A simple regression then could find a positive and significant relationship between the explanatory variable and the dependent variable, because both variables are independently trending upward at the same time, even if the two variables are not related. Moreover, the R^2 from an OLS regression has no meaning when data are

148. The Hausman Specification Test, which compares the fixed effects and random effects estimates, can be used for this purpose. If the zero correlation assumption is valid, the two sets of results should be similar. If the zero correlation assumption is invalid, then the random effects results (as well as least squares results) will be biased while the fixed effects results are still consistent, so that the two sets of results should diverge. In this case, the fixed effects model would yield more accurate results than either OLS or random effects. *See* Hausman, *supra* note 132.

149. WALTER ENDERS, APPLIED ECONOMETRIC TIME SERIES 181-184 (3d ed. 2010).

150. Roughly speaking, a series has a stochastic trend if the variability of the series increases over time. That is, if one forecasts future values of the series, the forecast uncertainty will increase with the forecast horizon.

not stationary. Many economic time series are thought to be nonstationary, such as a country's Gross Domestic Product or the price of a stock.[151]

The properties of many econometric estimators, such as least squares regression, are derived under the assumption that the data series defining the dependent and explanatory variables are not stochastically trending. If in fact the series are stochastically trending, a different set of econometric techniques should be employed.[152]

a. Tests for Stationarity

A number of different methods for testing for stochastic trends in a time series have been proposed.[153] These methods generally test for what econometricians call "unit roots."[154] Because of limitations in these unit root tests, one should view the results of the tests within the context of economic theory and the dynamics of the market. For example, suppose the time series of interest are price and marginal cost. In general, it would be economically implausible that one of these series is nonstationary while the other is stationary, because standard economic models predict that price is strongly influenced by marginal cost.[155] Now, suppose the results strongly reject the null hypothesis that marginal cost is nonstationary, but do not reject the null hypothesis that price is stationary. It would be reasonable to conclude from the results of the two tests—and the fact that a statistical rejection of one hypothesis is generally a stronger result than the inability to reject another hypothesis for the reasons discussed above—along with the implication of economic theory (suggesting both series must be either stationary or nonstationary) that both price and marginal cost are stationary in this case.[156] If so, then standard regression analysis can be used.

---

151. *Id.*

152. *Id.*

153. For a brief introduction to unit root tests, see GREENE, *supra* note 2, at 947-59; *see also* ENDERS, *supra* note 149, at 181-271 (providing a more extensive introduction to unit roots models and tests).

154. *See* GREENE, *supra* note 2, at 946.

155. *See id.* at 790-91. There are some situations in which price might be non-trending while cost was trending, or vice versa. For example, cost and a third explanatory variable that affects price could both be stochastically trending, but cointegrated (discussed further below). In that situation, a more complex econometric procedure must be used.

156. Similar considerations apply when one has panel data with long time series for each unit of observation. In many circumstances, it would be implausible for the price of one customer to be nonstationary while the

b. Modeling Market Dynamics with Stationary Time Series

Although the usual econometric techniques can be applied to stationary time series, particular issues may arise when attempting to model the dynamics of such series. The starting point for modeling those dynamics is studying the economics of the industry and examining why dynamic effects may arise in the first place. That is, why do the economic outcomes not adjust instantaneously to changes in economic conditions? This issue makes modeling only contemporaneous influences on the dependent variable inappropriate. There may be long-term contracts or other institutional features that slow the speed of adjustment, so understanding the economic context can be important in accurately modeling the dynamics.

Three types of dynamics may need to be modeled. First, dynamic effects may arise if the dependent variable is influenced by earlier values of itself. For example, where the dependent variable is price, the price in the current period can be affected by prices in prior periods if there are costs to firms changing their prices. These costs would discourage frequent price changes, so price changes would not occur as often and would likely be delayed until there was an accumulation of reasons to change price. Modeling this type of dynamic is often done by including lags (i.e., prior observations) of the dependent variable as additional explanatory variables.[157]

Second, dynamic effects may arise if the impact of a change in an explanatory variable plays out over two or more time periods. This can occur, for example, if long-term contracts allow for adjustments to price due to changes in costs, but only after a lag. Modeling these dynamics is often done by including not only the contemporaneous value of an explanatory variable, but also values of the variable from earlier time periods as additional explanatory variables.[158]

Finally, dynamic effects may arise if the error term includes unobserved factors that are themselves correlated over time. This induces *serial correlation* in the error term. The dynamics in this case may be

---

price of another customer is stationary. *See* J. Hlouskova & M. Wagner, *The Performance of Panel Unit Root and Stationarity Tests: Results From a Large Scale Simulation Study*, 25 ECONOMET. REV. 85 (2006).

157. WOOLDRIDGE, *supra* note 1, at 371-74.

158. *Id.*

limited to the error term, and can be addressed by modeling the form of the serial correlation.[159]

The economics of the industry can suggest that dynamics of one type or another are present, and the modeling should focus on that form of dynamics. As discussed above, if there is evidence that sales are made under contracts that adjust to cost changes with a lag, then lagged costs likely should be included as an explanatory variable for estimating the impact of cost on prices. In the case where there are reasons to believe dynamics of all three types may be present, it is important that the modeling of all three types be done jointly, rather than sequentially. There may be important interactions between the three types of dynamics that would be missed if the modeling were done sequentially.[160] Most basically, however, the key point is that if such dynamics are economically important they should be included in the model; otherwise results will suffer from a form of omitted variable bias due to the failure to account for the dynamic terms.

There is also an important question about the number of lags that should be included in the model. The optimal lag lengths then can be chosen by estimating the model using different combinations of lag lengths, and using model selection criteria such as the Schwartz criteria to choose between these models.[161]

Finally, it is important to recognize that specifications that include lags of the dependent variable as additional explanatory variables potentially complicate estimation of the model. In particular, such inclusion creates a correlation between the explanatory variables and the error term if the error term itself exhibits serial correlation.[162] Under this condition, least squares estimation of the model can lead to biased and inconsistent estimates, in which case other econometric methods should be used (as discussed below).

---

159. *See* GREENE, *supra* note 2, at 903-941. Note that, unlike failure to include appropriate lags of the dependent or independent variables, which introduces inconsistency into estimated coefficients, failure to account for the serial correlation in the error term only affects the efficiency of the estimates and the estimated standard errors, but does not lead to inconsistent parameter estimates.

160. WOOLDRIDGE, *supra* note 1, at 371-74.

161. GREENE, *supra* note 2, at 139-40.

162. *Id.* at 918-19.

c. Modeling Dynamics in Nonstationary Time Series

If the time series are stochastically trending, a different set of econometric techniques in general should be used to estimate the relationships between the series.[163] The simplest approach is to take the observed value of a variable in each time period observation "t" and subtract from it the value of the same variable in the previous time period "t-1." Doing this calculation for each observation creates a new time series data set of the changes in the variable from one period to the next, and this transformation is typically done for all of the variables involved in the regression.[164] This *first differences* approach frequently removes the nonstationarity from the data, although the transformed data should be tested to ensure that is in fact the case.[165] The approach, however, will also often remove some information about the relationship between the dependent variable of interest and the exogenous variables. Thus, goodness-of-fit measures such as R^2 and the measured significance of the statistical tests usually will be lower than regressions based on the untransformed data.[166] This loss of information also means that one should be cautious about using first differences approaches unless there is strong reason to believe the data exhibit nonstationarity.

More sophisticated econometric techniques can be applied if the time series in question are nonstationary, but are *cointegrated*.[167] Two (or more) nonstationary series are said to be cointegrated if a linear combination of the different series is itself stationary (or put differently, if the residual from a regression of one of the variables on the other is stationary).[168] For example, consider two time series. If one series plus a constant multiplied by the second series is stationary, the two series are cointegrated.

163. MADDALA & LAHIRI, *supra* note 34, 266-271, 551-578.

164. GREENE, *supra* note 2, at 959. Depending on the economics of the market and the nature of the alleged anticompetitive behavior, a dummy variable may not be transformed this way.

165. *Id.*

166. George C.S. Wang & Chaman L. Jain, Regression Analysis: Modeling & Forecasting 88 (2003).

167. If two series are nonstationary, but not cointegrated, regression techniques can still be applied. However, the estimated coefficients will have nonstandard sampling distributions. *See* Peter C. B. Phillips & Steven N. Durlauf, *Multiple Time Series Regression With Integrated Processes*, 53 REV. ECON. STUD. 473 (1986).

168. GREENE, *supra* note 2, at 959.

Econometric methods exist for testing whether two or more series are cointegrated.[169]

Cointegration can arise if economic forces prevent the nonstationary series from getting too far from each other, that is, if the series have a long-run equilibrium relationship. For example, again consider a price series and a marginal cost series. While price and cost may each be nonstationary and trend upward, the economic forces at play would tend to keep the gap between price and marginal cost (i.e., the margin) from becoming too large or too small, at least over a long enough time period. If the margin grew too large, entry would occur or customers would cut back their purchases and drive the margin back down. If the margin grew too small, exit would occur or customers would increase their purchases and the margin would increase again. Thus, the margin can be stationary even though price and cost are themselves nonstationary.

If the tests for cointegration indicate that the series are cointegrated, then the relationship between the series can be specified and estimated as a *vector error correction model* (VECM).[170] A VECM models both the cointegrating equations (the long-run relationships between the variables with common stochastic trends) and also the short-run dynamics of the residuals of the cointegrating equations.[171]

An important feature of these error correction models is that their estimated parameters typically have clear economic interpretations. The coefficients of the long-run economic relationships can be interpreted as measuring the long-run equilibrium relationships between the variables. So, for example, if the estimated error correction model relates prices to demand and cost measures, the coefficients should reflect how prices and demand and cost factors move together in the long-run. Thus, the economic reasonableness of the estimated model can be judged by comparing the parameter estimates to what economic theory would predict. If the parameter estimates are not consistent with economic theory, the estimates should generally be considered to be unreliable. Furthermore, the error correction model has an estimate of the speed with which the dependent variable adjusts toward this long-run equilibrium. Again, the reasonableness of this adjustment-speed parameter estimate can be judged based on economic knowledge.

---

169. *Id.* at 965-67. One of the most widely used of these tests is presented in Soren Johansen, *Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregression Models*, 59 ECONOMETRICA 1551 (1991).

170. GREENE, *supra* note 2, at 963-64.

171. *Id.*

d. Considerations in Building a Forecasting Model

Time series models may be used to make forecasts. For example, as discussed below, a time series model may be estimated using data from a competitive period, and then used to forecast but-for prices during a conspiracy period.[172] Such a forecasting model obviously will be useful only to the extent that it provides reliable forecasts. For example, the start and end of a conspiracy period may be characterized by sharp "turning points" in a price series. Accordingly, it is important that the forecasting model be able to predict turning points that occur for nonconspiracy reasons. Otherwise, turning points attributable to events unrelated to the conspiracy might be incorrectly attributed to the conspiracy.[173]

How does one tell if the forecasting model is reliable? Looking at R^2 is generally not a good way to assess a forecasting model. Models with lagged dependent variables or serially correlated errors often appear to track the historical data well and achieve a very high R^2. However, as discussed previously, such models may suffer from a form of "over-fitting."[174] That is, while they fit the historical data well, they fail to capture the true economic structure of the dynamic pricing process. The good fit may come from the inclusion of lagged dependent variables, not the inclusion other explanatory variables that predict changes in the dependent variable. Thus, when such models are used to forecast prices during an alleged conspiracy period, they may do a poor job forecasting turning points in the dependent variable.

The key criteria for assessing whether a regression model can be reliably used to make forecasts are succinctly summarized by Stock and Watson in their econometric text:

> [I]f we are to obtain reliable forecasts, the estimated regression must have good explanatory power, its coefficients must be estimated precisely, and it must be stable in the sense that the regression estimated on one set of data can be reliably used to make forecasts using other data. When a regression model is used for forecasting, a paramount concern is that the model is externally valid in the sense that it is stable and quantitatively applicable to the circumstances in which the forecast is made.[175]

---

172. *See* part F.1 of this chapter.
173. Baker & Rubinfeld, *supra* note 66, at 396.
174. *See* part D.6 of this chapter.
175. STOCK & WATSON, *supra* note 64, at 329.

Hence, an essential question that must be addressed in order to determine if a forecasting model can be used to forecast but-for prices is whether the economic model is stable between the estimation period and the anticompetitive period. The difficulty, of course, is that this is a question that involves comparing the model in the competitive period to what the model would have been in the anticompetitive period, absent the anticompetitive behavior. Thus, the forecasting approach does not avoid the fundamental problem of estimating "causal effects"—in fact, it makes the problem in some ways more complex because it requires that the stability of an entire equation be assessed.

One way to test a forecasting model is to see how well it predicts for some time period before the alleged anticompetitive act, assuming adequate data are available, or more generally to see if the forecasting model yields accurate predictions for periods assumed to be free of the allegedly anticompetitive act.[176] To implement such a test, one would divide the data from before the alleged anticompetitive act into two parts: one part is used to estimate the forecasting model, and the other part is used to assess the accuracy of the forecasts of the model. It is particularly useful to exclude from the part of the data used to estimate the model observations that contain a turning point (e.g., a substantial change in prices). If the forecasting model is reliable, then it should produce a reasonably accurate forecast for the part of the data not used in estimation. In particular, it should be able to predict any turning points. If the model does a poor job forecasting in this context, then it is unlikely to produce a reliable estimate of but-for prices during the period of the anticompetitive act.

Another way to test a forecasting model for stability is to perform Chow tests of stability during the competitive period. The logic is that if the econometric model is not stable during the competitive period, then there is a reasonable presumption that the model is unlikely to be stable when applied to the anticompetitive period. Furthermore, if there are two or more possible regimes during the competitive period, then it will require additional, careful economic analysis to determine whether any of these models are appropriate for the forecasting in the period during which the alleged anticompetitive conduct occurred. Of course, evidence of stability during the competitive period does not imply that the model estimated on the competitive period is appropriate for forecasting during the anticompetitive period, but evidence of instability during the competitive

176. KENNEDY, *supra* note 27, at 102.

period may create a presumption that forecasting models are unlikely to satisfy the required stability.

F. Before-During and Related Approaches to Damages

The before-during approach identifies the effect of the alleged anticompetitive conduct by using data from a period before the alleged conduct in combination with data from the period when the alleged conduct occurred.[177] Comparing the values of the dependent variable in the "before" period to its values in the "during" period may serve to identify the effect of the alleged conduct.[178]

To establish the extent, if any, of damages, it is generally useful to begin with a simple data analysis as a prelude to performing a more sophisticated analysis. For example, it may be useful to compare average prices before and after an alleged price-fixing agreement was implemented to explore the potential magnitude of a conspiracy-related overcharge. In industries in which prices tend to be relatively stable (e.g., vitamins, hydrogen peroxide), such price comparisons can often be done via a simple comparison of price levels during and outside the alleged cartel period. However, in industries in which prices are trending up or down (e.g., the high technology industries that have been the subject of several recent cartel investigations and in which prices are often falling rapidly), the more relevant question is generally whether, at the start or end of the alleged cartel, there is a break from the prevalent trend in prices.

In making this comparison, however, it is generally important to account for any significant differences in other economic factors between the before and during periods.[179] A basic comparison of average prices in different periods, however, does not hold constant other factors that may impact price. Thus, the initial analysis of averages is usefully supplemented with regression analysis.

---

177. *See* van Dijk & Verboven, *supra* note 19, at 2335-36.
178. *Id.* The before period may be considered the control period and the during period the treatment period. See the discussion in part I of this chapter.
179. *See* van Dijk & Verboven, *supra* note 19, at 2335-36.

1. "Dummy Variable" Models Versus "Prediction" Models[180]

Two methods are commonly considered to estimate the impact of alleged conduct on the outcome of interest: (1) dummy variable models; and (2) prediction models.[181] For simplicity, we compare the methods here in the context of a price fixing conspiracy, and as such in the estimation of price overcharges—though the ideas can be adapted to a broad range of outcomes and alleged conduct. Although both dummy variable and prediction models make use of data from a nonconspiracy period, they differ primarily in whether they also make use of data from the conspiracy period. The dummy variable model uses data from both the alleged conspiracy period and the non-conspiracy period to estimate the relationship between price, economic factors, and a dummy variable for the alleged conspiracy period, with the dummy variable measuring how much higher prices were in the alleged conspiracy period relative to the non-conspiracy period, after controlling for the other economic factors. The prediction model uses data from only the non-conspiracy period to estimate the relationship between price and economic factors. It then uses the estimated relationship as a basis to predict what prices would have been during the alleged conspiracy period but for the alleged conspiracy, with the gap between actual prices and the estimated but-for prices assumed to represent the effect of the price-fixing conspiracy.

For simplicity, assume that the best explanatory variable available for capturing the potentially anticompetitive effects of the price-fixing conspiracy is a dummy variable representing the time period when the alleged anticompetitive acts took place.[182] This variable will be 1 during

180. Material in this section draws heavily from Gustavo Bamberger, Dennis Carlton & Mark Israel, *Economic Analysis of Cartels* (Working Paper, drafted 2012).

181. Paul E. Godek, Time-Series Models for Estimating Economic Damages in Antitrust (and Other) Litigation: The Relative Merits of Predictive versus Dummy-Variable Approaches, 1 ANTITRUST CHRONICLE (2011); Philips & Durlauf, supra note 167.

182. Depending on the nature of the antitrust allegations, the theory of causation, and specific facts of the case, there can be variables other than a dummy that may better measure the anticompetitive behavior in estimating damages. *Id.* Assume, for example, prices can only increase when existing customer contracts expire over time. The impact of the alleged conspiracy would not increase prices to all customers immediately, so average price increases related to the alleged conspiracy would be expected to trend upward over time. There could be several ways to model this effect other than using a simple dummy variable for the alleged conspiracy period,

that period and 0 at other times (similar to the PERIOD variable discussed earlier). Other explanatory variables that affect the dependent variable may also be included in the regression. The coefficient on the dummy variable measures how much higher (or lower) the price would have been on average in the conspiracy period relative to the non-conspiracy period, controlling for other observable economic factors. Because the coefficient on the dummy variable in the regression model is an estimate of the effect of the conspiracy on the dependent variable, this type of model is often called a *dummy variable model*.[183]

This model typically assumes that the coefficients on the explanatory variables are the same in the conspiracy and non-conspiracy periods. This assumption may not always be valid. For example, price-fixing conspirators may react differently to changes in supply and demand factors than they would have had they not been conspiring. Thus, it is important to test whether prices react differently to the supply and demand factors in the conspiracy period than they did in the non-conspiracy period.

In more technical terms, it is important to test whether there has been a significant shift in the relationship between the explanatory variables and the dependent variable of interest.[184] In particular, one can estimate the same economic model over the conspiracy and non-conspiracy subsamples of the available data and test statistically whether the model's parameters are statistically significantly different in the two time periods. If the assumption of model stability is not rejected, then there is a statistical basis to rely on the results from the dummy variable model to estimate the

---

depending on the available data. *Id.* Under certain circumstances, one simple method to create a variable to approximate the conspiracy might be a "time trend" beginning at the period of the alleged conspiracy, with the value "1" in the first observation (*e.g.*, a month) of the conspiracy period, "2" in the second, "3" in the third, etc. If average prices were the appropriate dependent variable, then the coefficient on this trend variable would be interpreted as the average monthly price increase due to the conspiracy. The overcharge in any month during the conspiracy would be the corresponding time trend value multiplied by that coefficient. However, care must be used with this and any approach for creating a variable that measures the impact of an alleged anticompetitive act. For example, at some point all contracts will have been renegotiated, so it would be inappropriate to continue to predict price increases beyond a certain number of observations.

183. See James F. Nieberding, Estimating Overcharges in Antitrust Cases Using a Reduced-Form Approach: Methods and Issues, J. APPL. ECON. 361, 367-369 (2006).

184. GREENE, *supra* note 2, at 168-75.

cartel effect on price. If the stability assumption is rejected, the model may be unreliable.

An alternative to the dummy variable model is a *prediction model.*[185] The first step in implementing a prediction model is to estimate the reduced form equation on the data from the non-conspiracy period. Then, this equation is used to predict the prices that would have prevailed in the conspiracy period, given the values of the explanatory variables in that period.

The prediction model assumes that the coefficients on the explanatory variables that would have held during the conspiracy period absent the conspiracy are the same as those that actually hold in the non-conspiracy period. But, under certain conditions, this assumption may be invalid. For example, assume there was a change in the relationship between the explanatory variables and the variable of interest (e.g., price) in the conspiracy period that had nothing to do with the alleged anticompetitive conduct (e.g., entry or exit of firms from the market) and the prediction model does not take that change into account. Under these conditions, estimates using data solely from the non-conspiracy period are unlikely to accurately predict pricing but for an alleged conspiracy or other anticompetitive act.[186] Moreover, as discussed above, a prediction model should be tested where possible to check whether it accurately predicts turning points in the data.[187]

In addition, one also could estimate the model using data from the conspiracy period and then use the resulting estimates to predict back ("backcast") into the non-conspiracy period.[188] If the backcast is above the actual prices in the non-conspiracy period, that would suggest the price-setting process in the conspiracy period produced higher prices than the price-setting process in the non-conspiracy period. This finding would be consistent with the proposition that the alleged anticompetitive behavior

185. *See* MADDALA & LAHIRI, *supra* note 34, at § 4.7; Nieberding, *supra* note 183, at 367-69.

186. In this way, the prediction model may suffer from some of the same limitations related to the dummy variable model discussed above.

187. A dummy variable model can suffer from a similar problem of failing to fit turning points well, and thus incorrectly attributing to the conduct turning points that may have nothing to do with the alleged conduct. *See* WOOLDRIDGE, *supra* note 1, at 65-67.

188. Or, more generally, one could combine both data from before and after the period of the alleged anticompetitive act, using the combined before and after data to estimate the prediction model. Such before-during-after models are discussed in the next section.

in the conspiracy period increased prices, and the difference between the actual and predicted prices in the non-conspiracy period could form the basis for calculating an overcharge and damages. The backcast model relies on similar assumptions as the prediction model (in effect, it is predicting prices that would have prevailed had the conspiracy been operational in the non-conspiracy period).

If the results from the prediction and dummy-variable models are significantly different, one should assess carefully whether there are sufficient data and sufficient industry stability for the prediction model to be reliable. If the conditions required for a reliable prediction model are not met, then one should test whether the stability assumptions underlying the dummy-variable model are met, and if so, rely on the results from the dummy-variable model. Importantly, there may be situations in which neither model is appropriate, in which case the value of econometric evidence in assessing damages may be quite limited.

2. *Before-During-After Models*

The previous section discussed models that assume there is a single unaffected period (the period before the conspiracy). Sometimes, data might exist on several unaffected periods. For example, data may exist not only for the before period, but also for an after period. With more than one unaffected period, the question arises as to whether one should combine the two unaffected periods and, if not, then which of the two unaffected periods should be compared to the affected period.[189]

In general, the before and after periods should be combined, if appropriate, since that will result in a larger sample and more precise estimates.[190] However, it is appropriate to combine the periods only if the model is stable across them.[191] Although there are many ways that one might assess whether the model prediction model is stable across the two time periods (e.g., testing all or some of the coefficients for equality), the most informative test is likely to test whether a prediction model estimated using only the before period would yield the same overcharge estimates as a prediction model estimated using only the after period. If such equivalence of the before and after models cannot clearly be established, attempting to combine the before and after period to predict the path of

---

189. *See* van Dijk & Verboven, *supra* note 19, at 2336. Of course, it is important to ensure that the supposedly unaffected periods really are unaffected.

190. G.S. Maddala, Introduction to Econometrics 5, 69, 270 (3d ed. 2001); A.H. Studenmund, Using Econometrics: A Practical Guide 68 (4th ed. 2001).

191. Takeshi Amemiya, Advanced Econometrics 399-402 (1985).

prices "in the middle" (during the alleged anticompetitive act) is inadvisable, as predictions are likely to be unreliable.

A difficult decision arises when this test rejects combining the two periods into a single model and one must choose which unaffected period is the better basis of comparison for the affected period, assuming both are truly unaffected. Before addressing how to decide which unaffected period to use, first consider why two unaffected periods might be explained by different models. There are economic models of competition that suggest that periods of more intense competition ("price wars") will alternate with periods of less intense competition.[192] These models suggest that the relationship between price and the explanatory variables may change over time with the intensity of competition. In this case, a before period that was dominated by more intense competition might look different than an after period that was dominated by less intense competition, even though both periods were unaffected by the anticompetitive conduct at issue.[193] Alternatively, there may have been substantial differences in market conditions between the two periods (e.g., there may have been substantial new entry into the market), which are not adequately controlled for in the regression and therefore lead to economically relevant differences in the models

In the event that there are substantial differences between the before and the after periods, one needs to determine how to select which period to use. One way to answer this question is to try to model explicitly the switches between competitive regimes, and then use this model to predict which competitive regimes would have prevailed during the affected period in the but-for world. There are econometric techniques called *switching models* that can be used to do this.[194] Although not commonly done, using a switching model of this sort may also provide a realistic picture of competition in the affected period.

An alternative approach for choosing between the before and the after periods can involve using general market information and nonquantitative evidence from the case. For example, suppose that the after and during periods were characterized by strong demand for the product, while the before period was characterized by weak demand. These factors would suggest that the after period would provide a better basis of comparison for the affected period than would the before period.

192. See, e.g., Jerry Green & Robert Porter, Noncooperative Collusion Under Imperfect Price Information, 52 ECONOMETRICA 87 (1984).
193. *Id.*
194. AMEMIYA, *supra* note 191, at 399-402.

Another reasonable (albeit likely to result in lower damage estimates) approach may be to predict but-for prices for the affected period alternatively using the before- and after-period models, and then take the maximum of these two predictions. Notably, both of these predictions are consistent with competitive behavior, meaning that it may be reasonable to take the higher prediction as the price level consistent with competition with only prices above this higher level indicative of effects from the alleged conspiracy.

3. *Identifying Beginning and End Points of the Damages Period*

Recently, econometric methods have been developed that can be used to help in choosing beginning and end dates for allegedly unlawful conduct based on market data.[195] These methods involve estimating the point in time when an econometric model exhibits a *structural change*, or shift in the model parameters. In the case of alleged anticompetitive conduct that had an effect on market outcomes, a structural shift in the econometric model should have occurred when the effect of the conduct began and a second structural shift should have occurred when the effect of the conduct ended. Thus, for example, identifying when there were structural shifts in the regression model can provide useful evidence concerning the beginning and end of overcharges from a price-fixing conspiracy.[196]

However, it is statistically inappropriate to identify a structural break in a data series and use that break to define a conspiracy period unless one properly adjusts the statistical inference for how the beginning and end points are chosen. For example, it would not be appropriate to identify a price spike, define the beginning of a conspiracy period from that spike, and then "test" whether prices were higher in the period beginning with the spike than before the price increase. Indeed, the literature has shown that approaches that search for beginning or end dates without using the appropriate statistical tests are very likely to find statistical evidence of "damages" even in cases where there was no anticompetitive conduct.[197]

195. Bai & Perron, *supra* note 142. *See* Dennis Carlton & Gregory K. Leonard, *Correcting the Bias When Damage Periods Are Chosen to Coincide With Price Declines*, 2004 COLUM. BUS. L. REV. 304 (2004), for a discussion of the application of these methods to price-fixing cases.

196. Carlton & Leonard, *supra* note 195.

197. *Id.*

G. Benchmark and Difference-in-Differences Approaches to Damages

Two additional methods are sometimes used to estimate the impact of alleged conduct on the outcome of interest: (1) benchmark models; and (2) difference-in-differences models. Both approaches to damages are based on the same basic principles and econometric tools as the before-during-after approach.

The benchmark analysis compares the outcome of interest (e.g., price) in the affected market to that same outcome in an unaffected benchmark market.[198] The benchmark market (or markets) can be a different geographic area or a different end-use market, although when using a different end-use market it can be particularly challenging to control sufficiently for all of the differences between the markets. The key is that the benchmark markets, however chosen, were unaffected by the alleged anticompetitive conduct. Further, the benchmark and affected markets should be characterized by sufficiently comparable economic conditions (at least after controlling for observable factors) such that prices in those markets would have been identical had there been no anticompetitive behavior.[199] The benchmark approach may be particularly useful if there are inadequate data for the before period, so that a before-during approach cannot reliably be performed.

The difference-in-differences model compares the differences in outcomes before and after the alleged anticompetitive conduct across the benchmark and affected markets. Under the assumption that the differences would have been stable, controlling for observable factors, but for the alleged anticompetitive conduct, the difference-in-differences provides an alternative approach to estimating damages. (Said differently, the difference-in-differences model requires an assumption that the market outcome would have changed in the same way in both the benchmark and affected markets, after controlling for observable factors, absent the alleged anticompetitive conduct.) The difference-in-differences approach can be implemented using a regression that takes into account the various factors that differ between markets and time periods.

The difference-in-differences approach requires data for the market of interest and the benchmark market that cover both a period of time that is unaffected by the alleged anticompetitive act and the period of time when the market of interest allegedly was affected. Thus, the difference-in-

198. van Dijk & Verboven, *supra* note 19, at 2336.
199. *Id.*

differences approach combines the before-during and benchmark approaches.[200]

To take a simple example, suppose that Market A is presumed to be affected by a conspiracy and Market B is not. Suppose that prices in the non-conspiracy period in Market A are $15 and in Market B are $10. Further suppose that the prices in the conspiracy period in Market A are $40 and in Market B are $20. A basic before-during analysis, without controlling for anything else, would conclude that prices in Market A are $25 higher in the during period than the benchmark period. However, this estimate is likely to overstate the effect because prices were rising in general (e.g., in Market B). Similarly, a benchmark approach that compared Market A to Market B in the conspiracy period without controlling for anything else would conclude that prices are $20 higher in Market A than Market B. However, this estimate is also likely to overstate the effect of the conspiracy because it fails to take into account that prices in Market A are generally higher than in Market B (e.g., during the benchmark period). The difference-in-differences approach combines these two approaches by recognizing that the difference between prices in Market A and Market B was $5 in the benchmark period and $20 in the during period. Thus, it would conclude that the effective price difference was really $15.[201]

H. Summary of Key Points in Creating and Evaluating Econometric Models of Damages

As can be surmised from the previous sections, a number of considerations should be taken into account when building or evaluating an econometric model used to estimate damages.[202] Model specification and estimation involve many choices: explanatory variables, functional form, dynamic specification in a time series, estimation technique, and which of the basic approaches to damages to use. In many cases, statistical tests can help guide these choices. However, the starting point for any econometric model is economic theory and knowledge of the industry, along with confidence, that there are sufficiently reliable data to perform an econometric analysis.

---

200. WOOLDRIDGE, *supra* note 1, at 148.

201. A similar analysis could be undertaken based on percentage rather than level differences.

202. Much of the discussion here and elsewhere in this chapter can be applied to any econometric analysis. *See, e.g.*, WOOLDRIDGE, *supra* note 1; GREENE, *supra* note 2.

Economic theory provides guidance as to the general types of explanatory variables that should be included in the model. For example, for an econometric model of price, economic theory suggests that variables representing cost are likely the most important control variables, with variables capturing demand and competitive factors also potentially important types of explanatory variables.[203] Knowledge of the industry helps determine with more specificity which variables to include. For example, if the industry in question were plastics, the price of petroleum (an important input to plastics) likely would be a relevant cost variable and production levels in downstream industries that use plastics likely would be relevant demand variables. In general, to avoid "data mining" or "fishing," it is important to limit explanatory variables to those that are highly economically relevant to the industry in question.

One must be able to obtain enough sufficiently reliable data to implement any economically sound econometric model. Econometric analysis based on very few data points in general will be less reliable than econometric analysis based on a larger number of data points. There may also be issues with how accurate the data are, and whether they represent an unbiased sample of what has occurred in the market. Accordingly, the data need to be checked for accuracy and any potential selection bias. Finally, in many instances the data will need to be "cleaned" to some extent, removing obviously incorrect observations. For example, many damages analyses use data from a company's sales database. Care must be taken with such data to ensure that all significant discounts have been accounted for, and to weed out any obvious errors.[204] In some circumstances, limitations on the availability or quality of data may make it impossible to conduct a meaningful econometric analysis.

Often economic theory and industry knowledge might suggest a large number of potential explanatory variables. While there is a benefit to parsimony (namely potentially increased precision of the estimators), there may be a greater concern with omitting a potentially important explanatory variable that can induce omitted variable bias. Again, however, the set of explanatory variables considered for inclusion should generally be limited to those that are highly economically relevant for the industry in question and testing should be done to make sure results are not overly sensitive to small changes in the set of explanatory variables included in the regression.

---

203. van Dijk & Verboven, *supra* note 19, at 2335.

204. For example, frequently these transactions data sets include credits for returned goods, etc. It is important to use net revenue and net sales to calculate a measure of the price actually paid.

The relationship between the dependent variables and the explanatory variables can be expressed in various functional forms of the underlying variables, and this is another important aspect of model specification.[205] For example, the dependent variable might be the unadjusted price of the product at issue and an exogenous variable could be costs. This model would provide an estimate of how much price would go up for a dollar increase in cost. However, there may be good theoretical or industry-specific reasons to believe that the relationship of price to cost is better measured in percentage change terms for each variable. If so, then the functional form of the estimating equation may be better specified using the logarithms of price and cost as variables. Statistical tests can help determine which functional form is most appropriate for the situation at hand.[206]

In a time series context, specification of the dynamics is important. Again, the particular specification might be guided by industry knowledge and statistical tests, including those that can help make choices about (for example) the lag lengths.[207]

Care must be taken to specify correctly the structure of the error term in the estimation of the econometric model and to use estimates of the standard errors that are robust to deviations from the assumption of independent, identically distributed error terms. Only with correct standard errors will the resulting statistical inference be valid. The correct method for calculating standard errors can be determined based on the outcome of statistical tests, and in general such tests should be performed.[208]

The choice between reduced form and structural models can be of great importance. A reduced form model typically will provide the most efficient and simplest model for damage estimation, but in some instances estimating a structural model will be a superior approach.[209]

The decision regarding which damages approach to employ also will depend on the available data and the nature of the market. Are there sufficient data to implement a before-during or before-during-after model? Is there a good benchmark? Can a difference-in-differences approach be used? In general, if reliable data are available from outside the period of

---

205. *See* part D.1 of this chapter.
206. *Id.*
207. *See* part E.3 of this chapter.
208. *See* part D.5 of this chapter.
209. *See* part D.3 of this chapter.

the alleged anticompetitive act, then techniques that take advantage of those data will likely be the most reliable.[210]

Deciding whether to use a dummy variable or prediction model can also be important. As noted above, if the results from the predictive and dummy-variable models are significantly different from one another, then one should first assess carefully whether there are sufficient data and sufficient industry stability for each of these models to be reliable, being mindful that there may be situations in which the conditions for neither the predictive nor dummy-variable model are met.

Once the econometric model is specified and estimated, one must subject the results to a series of sanity checks and another series of statistical tests to assess the reliability and robustness of the estimates.[211] The starting place is typically examining the sign and statistical significance of the coefficients. If, for example, the coefficient on cost in a price regression is negative and statistically significantly different from zero (or very small and/or not significantly different from zero), this would cast doubt on the reliability of the model, since one would normally expect cost to have a positive effect on price. When performing this sort of check, however, one must take care that there is a clear interpretation to the coefficients; there may not be one in a reduced form model, for instance.[212]

When there is reason to question a model's specification, a variety of specification tests can and should be employed. For example, as discussed above, a model may impose potentially questionable restrictions on the coefficients, such as that the coefficients are the same during two time periods. Specifically, an econometric model of pricing at the customer level might assume that all customers have the same coefficients on the cost and demand variables. However, there may be reason to think that customers have different responses to cost and demand conditions, so there could be different coefficients on the explanatory variables across customers, including the variable measuring damages. In this case, it may be necessary to estimate the model for different groups of customers, or otherwise to employ a more complex model that takes these differences into account.[213] Similarly, a time series model might assume that the coefficients are the same over the course of the entire time period, but there may be evidence that the structure of the model changed at some point in time. A Chow test (or a more sophisticated test for structural breaks) can

210. *See* parts F and G of this chapter.
211. *See* parts C.4 and D.6 of this chapter.
212. Baker & Rubinfeld, *supra* note 66, at 392.
213. *See, e.g.*, Johnson & Leonard, *supra* note 48, at 351-52.

be used to test whether such a change occurred.[214] If so, the appropriate model should adequately control for such a change.

There may be reasons to question the crucial assumption that the error term of the regression is uncorrelated with the explanatory variables. This assumption often can be tested using a form of the Hausman specification test. For example, if we are concerned that an explanatory variable might be endogenous (which would bias OLS regression results), then the Hausman specification test comparing the OLS results to the IV results can be used to check for this problem.[215] If such a problem is found to exist, then more complex econometric techniques can be used to correct for the bias.[216]

Finally, virtually any regression model eventually will fail one or more tests if enough tests and specifications are run, even if nothing is wrong with the model. However, failure of a test should be taken seriously, and a model should be rejected when it fails a test of a *critical* assumption, or if it fails a large number of the specification tests to which it is subjected.[217]

More important than whether the model fails one of many possible statistical results, is that the results are fairly robust to a range of reasonable, small tweaks to the model. If slight changes to the functional form of the model, the precise set of control variables used to capture relevant economic factors, or the number of lags of the dependent or explanatory variables lead to large changes in estimated damages, then the results from the model should generally be treated with substantial skepticism.

I. Case Study—The Application of the Before-During and Benchmark Approaches

Econometric techniques can be applied in a variety of ways to estimate the impact of the alleged anticompetitive conduct on certain economic outcomes, such as estimating lost sales or price changes due to the alleged conduct. Some of the econometric techniques and tests discussed above and their applications are illustrated using the before-during and benchmark approaches to estimate the price effects of alleged collusive

---

214. The Chow test is a statistical and econometric test of whether the coefficients in two linear regressions on different data sets are equal. *See* Gregory C. Chow, *Tests of Equality Between Sets of Coefficients in Two Linear Regressions*, 28 ECONOMETRICA 591 (1960).

215. Hausman, *supra* note 132.

216. WOOLDRIDGE, *supra* note 1, at 89-121.

217. KENNEDY, *supra* note 27, at 76-79.

behavior by analyzing actual data relating to Southern Powder River Basin (SPRB) coal.

1. Background

Virtually all coal produced in the United States is used for electric power generation.[218] About one-third of the coal now comes from the SPRB, which contains vast amounts of low-sulfur coal available in low-cost surface mines. SPRB coal is relatively homogeneous, with two main types: 8400 Btu/lb. coal and 8800 Btu/lb. coal. SPRB coal mining began with the 8400 Btu coal, which was closer to the surface and easier to mine. During the 1970s, many coal-powered electric generation units were constructed to burn this coal. Virtually all of these units also can burn the higher-quality 8800 Btu coal. As sulfur emissions regulations became more stringent, older electric generation units were converted from burning 12,000 Btu bituminous coals to burning 8800 Btu SPRB coals. Because these plants were already suffering capacity losses from burning lower Btu SPRB coals, they generally did not burn 8400 Btu coal.

During 1999 and 2000, the three leading SPRB coal producers conducted what economists may describe as "cheap talk." Cheap talk refers to communications between oligopolists about future intentions. Such talk can, in some circumstances, facilitate understanding of each other's goals and intentions, which can lead to collusive production and pricing.[219] The statements were of the nature that they would reduce production "until prices improved." The largest producer, Peabody Energy, closed a mine in 1999. The third largest producer, Arch Coal, Inc., closed another mine in 2000, and several other planned expansions were cancelled. The result was a spot price increase in 2001. In the aftermath of the capacity reductions, prices settled about 30 percent above the level before the shut-downs. The DOJ opened an investigation in July 2001 but ultimately took no action against the SPRB producers.

---

218. U.S. ENERGY INFORMATION ADMINISTRATION, ANNUAL COAL REPORT 2007 9 (2009), *available at* http://www.eia.doe.gov/cneaf/coal/page/acr/acr.pdf.

219. *See* MICHAEL D. WHINSTON, LECTURES ON ANTITRUST ECONOMICS 20 (2006). Whether such behavior constitutes an unlawful price-fixing "agreement" is a legal question.

2. *Economic Model, Data and Choice of Explanatory Variables*

For the purpose of this case study, the simplest economic model hypothesizes that if the alleged collusion had an effect, then prices of SPRB coal during the damages period should be higher than before the damages period after controlling for other supply and demand influences. As will generally be the case, the most straightforward econometric models for estimating damages use a reduced form, rather than a structural form,[220] but the model still needs to include the major supply and demand influences on price that would be included in a structural model. Here, we assume a linear relationship between the relevant market forces and SPRB prices.[221] (A full analysis would investigate whether the results were robust to alternative specifications of the relationship of these factors to price, such as in percentage change terms.)

The price of SPRB coal used in the analysis is the spot price reported by the utilities that purchased the SPRB coal.[222] The utilities reported the spot price at the time of delivery instead of at the time of actual sales. As a result, there is a lag between the reported spot price and the spot price at the time of actual sales that ranges from one quarter to one year. Therefore, although the alleged collusion occurred in the first quarter of 2001, the reported spot price did not increase until late 2001. This case study defines the before-collusion period as before November 2001 and the during-collusion period as between November 2001 and May 2005.

Before one can engage in econometric analysis of the data, it is important to review the data and check whether there are potential coding

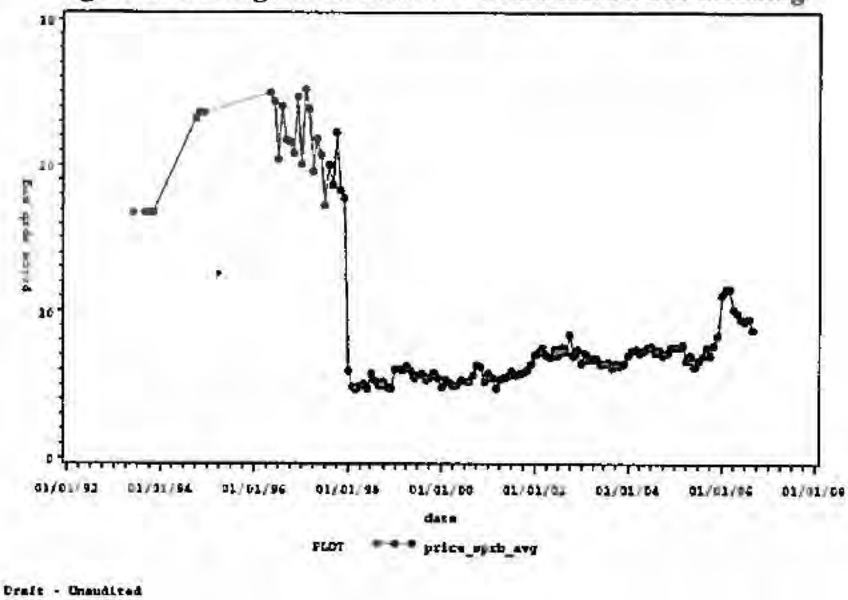
220. See part D.3 of this chapter for a discussion of structural form models versus reduced form models.

221. The linearity assumption is, computationally, the easiest assumption one can make. It assumes that a change in one of the relevant market forces has a proportional effect on the outcome variable, price in this case. Further, the estimated coefficients are themselves the marginal effects of changes in the explanatory variables on the outcome. By contrast, if there were non-linearities between the relevant market forces and the outcome variable, a change in one relevant market force would have a disproportional effect on the outcome, and the marginal effects of changes in the explanatory variables would have to be computed with the estimated coefficients.

222. The data source is FERC Form-423; Energy Velocity. Regulated utilities report the delivered price of their fuels in FERC Form-423. Energy Velocity estimates transportations costs and subtracts these from the delivered costs to provide a comparable FOB mine price for each utility. The average of these monthly spot FOB prices provides data for our dependent variable.

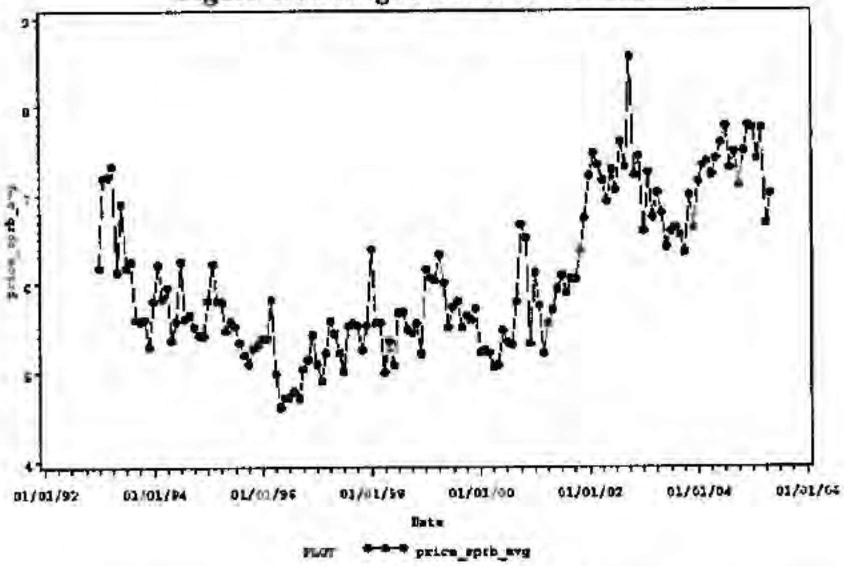
errors or other issues in the data that deserve further investigation. As discussed in section A, data do not have to be perfectly accurate to be useful in econometric analysis, but it is important to ensure that the data are reasonably reliable. To illustrate the importance of initial data review before running any regression analysis, the original data on the average SPRB coal price are plotted in Figure 3.

Figure 3: Average Price of SPRB Coal before Data Cleaning



As can be seen clearly from Figure 3, something is not right with the SPRB coal price data before 1998. Further review of the data reveals that before 1998, the transportation cost from the coal mines to the utilities was in the column for the spot price and vice versa. Correcting these data and conducting some other cleaning of the data, such as estimating the spot prices that were missing, results in data on the corrected average spot price for SPRB coal, presented in Figure 4.

Figure 4: Average Price of SPRB Coal



These prices represent the prices of SPRB coal delivered from January 1993 to May 2005. Through most of the 1990s, these prices display significant volatility, but they did not have any uniform trend up or down. Beginning in November 2001, prices began increasing from average levels of around $5.60/ton in the before period to over $7.00/ton in the alleged conspiracy period. The question is whether this increase in prices was caused by changes in demand, changes in costs, or changes in the behavior of mine operators.

To take into account the major demand and supply factors that potentially could affect the prices of the SPRB coal, data on these factors were obtained from public sources. Supply data include quality measures of SPRB coal (BTU content, ash percentage, and sulfur content); productivity of the mines that produce the SPRB coal as measured by the average tons of SPRB coal produced per person day; and costs of fuels that are used in the production of SPRB coal, including diesel fuel cost and electricity cost.[223] Demand was measured using data on total electric power production.[224]

---

223. Time series data on capacity for SPRB coal production were not available. If the data were available, one may want to consider capacity in the analysis as well since it may affect the supply of SPRB coal.

224. The spot prices of SPRB coal were reported by the utilities at the time of delivery instead of at the time of sales, so there is a lag between the time

3. Assumption of Liability and Causation

In performing a damage analysis, experts usually assume that the plaintiff has been found liable for the alleged anticompetitive conduct. Accordingly, this case study assumes that SPRB coal producers engaged in collusive behavior from November 2001 to May 2005. As discussed in section C, however, even assuming collusion did occur, one still needs to study whether the observed price increase was at least in part due to other demand or supply changes or whether it could be attributed to the collusive behavior. To help answer this question and estimate the price increase that resulted from the assumed collusion, the dummy variable model and the prediction model under the before-during approach are used to isolate the price effects of the alleged collusive behavior.

4. Stationarity of the Time Series

As discussed in section D, when using time series data to conduct regression analysis, one should test whether the time series are stationary. If they are not stationary, different econometric techniques should be employed. To study whether nonstationarity is an issue in the time series data, Dickey-Fuller/Augmented Dickey-Fuller unit root tests were conducted[225] on the SPRB coal price data and certain independent variables.

The unit root test results indicate that the SPRB coal price data are stationary while some of the independent variables may be nonstationary, such as the diesel price series. However, as pointed out in section D, many unit root tests have a low ability to reject nonstationarity. Therefore, one should view the results of the tests within the context of economic theory. In this case, if the dependent variable, SPRB coal price, is stationary, it is unlikely that the demand or supply factors that interact with each other to determine the SPRB coal price are nonstationary. If these demand or

---

when the prices were reported and the time when the prices were determined. This delay in reporting is not an issue for other independent variables, such as diesel cost, electricity cost, electricity production and SPRB coal productivity measurement. Therefore, in theory, one should lag these independent variables for certain periods and include these lagged variables as independent variables in the regression as well. However, this undertaking is beyond the scope of this illustration.

225. *See* MADDALA & LAHIRI, *supra* note 34, at 548-52. In the Dickey-Fuller/Augmented Dickey-Fuller test, the null hypothesis is that the time series is nonstationary. *Id.* For a brief introduction to stationarity tests, see MADDALA & LAHIRI, *supra* note 34, ch. 14.

supply factors were nonstationary, they likely would drive the SPRB coal price to be nonstationary as well.

To further investigate the nonstationarity issue, additional tests were conducted on the residuals from an OLS regression where the SPRB coal price is the dependent variable and other demand and supply factors along with a collusion dummy variable are included as the explanatory variables.[226] These stationarity tests indicate that the residuals from the OLS regression are stationary. Since the SPRB coal price is stationary, this result again suggests that nonstationarity of other explanatory variables is not an issue in the case study. Thus, an OLS regression analysis was conducted to assess how the assumed collusion affected the SPRB coal price.

5. *The Before-During Approach with the Dummy Variable Model*

The dummy variable model under the before-during approach can be used to measure how the assumed collusion changed the price of SPRB coal.[227]

---

226. The OLS regression model specification is discussed in more detail in part C.2 of this chapter. In conducting the Dickey-Fuller/Augmented Dickey-Fuller tests of the OLS regression residuals, one should use the Engle-Granger critical values instead of the standard Dickey-Fuller critical values because the Dickey-Fuller test of the residuals is a co-integration test. For more details on the Engle-Granger critical values, see ENDERS, *supra* note 149, at 373-82. However, the exact Engle-Granger critical values for these tests were not available because the number of independent variables in our regression exceeds those listed in the table of Engle-Granger critical values in Enders. Therefore, the KPSS test was also performed. In the KPSS test, the null hypothesis is that the time series is stationary. See MADDALA & LAHIRI, *supra* note 34, ch. 14, for more details.

227. The dummy variable is constructed by creating a variable equal to 0 before November 2001 (before period) and equal to 1 from November 2001 onward (during period). Together with other explanatory variables, the model is specified as $SPRB\_price_t = \alpha + \beta_1 \cdot Btu_t + \beta_2 \cdot Ash_t + \beta_3 \cdot lbsSO2_t + \beta_4 \cdot ton\_prod_t + \beta_5 \cdot diesel\_price_t + \beta_6 \cdot electricity\_price_t + \beta_7 \cdot electricity\_production_t + \gamma \cdot dummy_t + \varepsilon_t$, where Btu_t is the Btu content of the SPRB coal, Ash_t is the ash percentage of the SPRB coal, $lbsSO2_t$ is the pounds of sulfur per million Btu of energy in the coal, ton_prod_t is the average tons of SPRB coal produced per person day, $diesel\_price_t$ is the price of diesel, $electricity\ price_t$ is the price of electricity, $electricity\_production_t$ is the production of electricity, $dummy_t$ is the dummy variable, and ε_t is the error term.

If the assumed collusion led to an increase in the price of SPRB coal, the coefficient on the dummy variable should be positive and statistically significant after controlling for the demand and supply factors that affect the price of SPRB coal. As discussed in part D of this chapter, valid hypothesis testing requires consistent estimates of the standard error of the coefficient estimates. Tests for serial correlation indicate that the error terms in the regression are serially correlated. To estimate the standard errors consistently, the Newey-West procedure was used to obtain the heteroskedasticity and autocorrelation consistent Newey-West standard errors.[228] The result of the OLS regression with the Newey-West standard errors is presented in Table 1.[229]

Table 1: Regression Results for Dummy Variable Method

Dependent Variable:	Average Price of SPRB Coal
R-squared:	0.7132
Adj R-squared:	0.6968
Number of Obs:	149

	Coefficient	Newey-West Standard Error	t-Statistic	P>|t|
Dummy variable for the alleged collusion period	1.2362	0.1798	6.8800	0.0000
Btu content of SPRB coal	-0.0013	0.0011	-1.1300	0.2590
Ash percentage of the SPRB coal	-1.0727	0.7633	-1.4100	0.1620
Pounds of sulfur per million Btu	-1.1665	1.5701	-0.7400	
Average tons of SPRB coal produced per person day	0.0015	0.0024	0.6200	0.5330
Price of diesel	0.0978	0.3283	0.3000	0.7660
Price of electricity	0.0497	0.0483	0.2400	0.8080
Production of electricity	4.33E-10	1.78E-09	0.2400	0.8080
Constant	20.7707	11.0578	1.8800	0.0620

228. *See* Jeffrey M. Wooldridge, Introductory Econometrics—A Modern Approach 410-13 (2d ed. 2003).

229. Obtaining the Newey-West standard error requires determining the maximum order of lag in advance. This illustration follows the practice of using the smallest integer greater than or equal to T1/4 where T is the sample size. *See* GREENE, *supra* note 2, at 465.

As can be seen from Table 1, the variables that control for supply and demand changes are of the expected signs. Individually they are not statistically significant from zero at the 5 percent level, although they are statistically significant when taken together.[230] The coefficient on the dummy variable is positive and statistically significant. The regression result indicates that on average, SPRB coal prices were $1.24/ton (or about 22 percent) higher in the period with the unlawful conduct, compared to the before period, holding constant the exogenous variables.

As discussed in part F of this chapter, the dummy variable model assumes that the relationship between the dependent variable and the explanatory variables in the during period is the same as in the before period. To test this assumption, interaction terms involving the dummy variable and each of the other explanatory variables except the intercept term were added as explanatory variables in the model specification as well.[231] The null hypothesis that all of the coefficients on the interaction terms are jointly equal to zero was tested. If this hypothesis is not rejected, then the data are consistent with the view that the relationship between the price of SPRB coal and the demand and supply factors did not change from the before period to the during period. If the null hypothesis is rejected, it suggests that the relationship has changed, either due to the alleged behavior or because of other reasons.[232]

As shown in Table 2, the test rejected the null hypothesis that the relationship between the price of the SPRB coal and the demand and supply factors remained the same in the during period.[233] Hence, a possibility that must be investigated is that the restrictions on the coefficients of the demand and supply factors imposed by the dummy variable model may have substantially affected the model's estimate of the effect of the alleged antitrust violation. Accordingly, the next section uses

---

230. In a more detailed analysis, other measures of supply and demand might be considered or the estimating equation might be modified to better take into account more specific demand and supply considerations for each of the utilities.

231. These interaction terms are equal to zero in the pre-collusion period and equal to the respective values of the original explanatory variables during the collusion period. *See* MADDALA & LAHIRI, *supra* note 34, at 307-310.

232. *See* part E.1 of this chapter.

233. To implement the joint F test correctly, again, the Newey-West covariance matrix is used. *See* WOOLDRIDGE, *supra* note 228.

the prediction model to assess the impact of the assumed collusion on the price of SPRB coal.[234]

---

234. When one fully interacts the dummy variable with other explanatory variables, one in fact is assuming that the regression equations in the before period and in the during period are completely different. *See* MADDALA & LAHIRI, *supra* note 34, at 309. For example, the regression equation in the during period can have smaller intercept terms, hence a negative coefficient on the dummy variable, but higher slope terms. As a result, one cannot interpret the coefficient on the dummy variable as the measurement of the average impact of the alleged collusion anymore.

Table 2: Regression Results—Testing for Relationship Change in the Dummy Variable Method

Dependent Variable:	Average Price of SPRB Coal
R-squared:	0.7417
Adj R-squared:	0.7126
Number of Obs:	149

	Coefficient	Newey-West Standard Error	t-Statistic	P>|t|
Dummy variable for the alleged collusion period	-57.0152	23.4560	-2.4300	0.0160
Btu content of SPRB coal	-0.0021	0.0016	-1.3000	0.1940
Ash percentage of the SPRB coal	-1.9244	0.8683	-2.2200	0.0280
Pounds of sulfur per million Btu	0.9237	1.7494	0.5300	0.5980
Average tons of SPRB coal produced per person day	0.0038	0.0025	1.5300	0.1280
Price of diesel	-0.0652	0.4501	-0.1400	0.8850
Price of electricity	0.0296	0.0860	0.3400	0.7310
Production of electricity	9.85E-10	2.32E-09	0.4300	0.6710
Interaction term for Btu content of SPRB coal	0.0056	0.0025	2.2600	0.0260
Interaction term for ash percentage of the SPRB coal	2.5203	1.2294	2.0500	0.0420
Interaction term for pounds of sulfur per million Btu	-5.2794	2.4693	-2.1400	0.0340
Interaction term for average tons of SPRB coal produced per person day	0.0038	0.0081	0.4700	0.6370
Interaction term for price of diesel	0.5961	0.5924	1.0100	0.3160
Interaction term for price of electricity	-0.0406	0.1036	-0.3900	0.6960
Interaction term for production of electricity	-6.07E-10	2.96E-09	-0.2100	0.8380
Constant	29.8307	14.1026	2.1200	0.0360
Joint F test for all interaction terms are equal to zero	F(7, 133) = 3.09		Prob>F = 0.0048	

6. The Before-During Approach with the Prediction Model

The prediction model assumes that the relationship between the price of SPRB coal and the demand and supply factors in the actual world before the alleged anticompetitive behavior would have remained the same but for the alleged behavior in the during period.[235] If the relationship change found in the last section is entirely due to the collusive behavior, then the prediction model still would be valid. It is important to note, however, that the prediction model assumes *all* changes in the observed relationships between the before period and the during period are due to the collusive behavior; hence, in cases where the industry is changing rapidly, prediction models may confound other changes with changes due to collusive behavior and thus may not be reliable.

To use the prediction model to estimate the impact of the assumed collusion on the price of SPRB coal, an OLS regression similar to the dummy variable model was run, but using the data in the before collusion period only and excluding the dummy variable from the explanatory variables. The result of this regression is presented in Table 3.

Table 3: Regression Results for Prediction Model

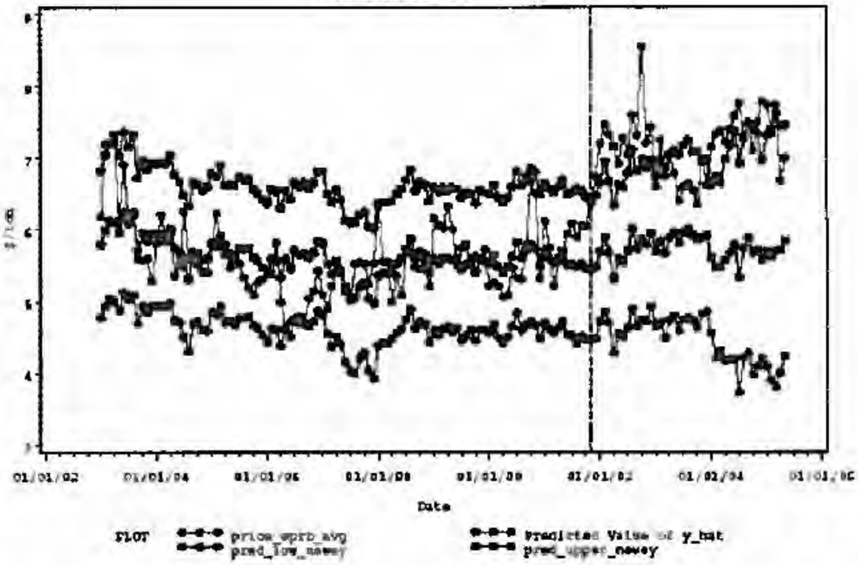
Dependent Variable:	Average Price of SPRB Coal
R-squared:	0.2104
Adj R-squared:	0.1540
Number of Obs:	106

	Coefficient	Newey-West Standard Error	t-Statistic	P>|t|
Btu content of SPRB coal	-0.0021	0.0015	-1.3300	0.1870
Ash percentage of the SPRB coal	-1.9244	0.8532	-2.2600	0.0260
Pounds of sulfur per million Btu	0.9237	1.7189	0.5400	0.5920
Average tons of SPRB coal produced per person day	0.0038	0.0025	1.5600	0.1220
Price of diesel	-0.0652	0.4423	-0.1500	0.8830
Price of electricity	0.0296	0.0845	0.3500	0.7270
Production of electricity	9.85E-10	2.28E-09	0.4300	0.6660
Constant	29.8307	13.8571	2.1500	0.0340

235. See the discussion in part F of this chapter.

Given the coefficients from the before period OLS regression, the next step is to predict the price of SPRB coal in the but-for world without collusion, using the data on the explanatory variables in the during period. Figure 5 shows the actual price of the SPRB coal, the predicted price of the SPRB coal based on relationships estimated in the before period, and the lower and upper bound of the 95 percent prediction interval for the predicted price of the SPRB coal.[236]

Figure 5: Average vs. Predicted Price of SPRB Coal—Newey Standard Error



As can be seen in Figure 5, in every month in the during period, the actual price of the SPRB coal is higher than its predicted price. Moreover, in 28 of the 43 months in the during period, the actual price of the SPRB coal is higher than the upper bound of the 95 percent prediction interval for the predicted price of the SPRB coal. The average predicted price of the SPRB coal for the during period is $5.72/ton, while the average actual

236. See GREENE, *supra* note 2, at § 6.6, for detailed discussion on how to construct a confidence interval for the predicted values. Constructing a confidence interval for the predicted value requires estimating the variances of the predicted value and the error term. The illustration uses the Newey-West standard errors in estimating the variance of the predicted value and uses the residuals from the OLS regression to estimate the variance of the error term.

price of the SPRB coal is $7.16/ton in the during period. Therefore, according to the prediction model, the assumed collusion resulted in a $1.66/ton (or about a 25 percent) increase of the SPRB coal price.

In summary, the prediction model results suggest a slightly larger impact of the assumed collusion on the price of the SPRB coal compared to the $1.24/ton (or 22 percent) increase of the SPRB coal price based on the dummy variable model. But critically, the results from both of these approaches are consistent with the alleged behavior causing higher prices. Hence, the result that prices were higher during the collusive period is robust across markedly different econometric methods, greatly increasing the confidence one should have in this finding.

Given that the test shown in Table 2 above rejected the hypothesis that the relationship between the price of the SPRB coal and the demand and supply factors were the same in the before and during periods, the prediction approach would appear to be the more reliable quantification of the impact of the hypothesized conspiracy. However, to determine if the prediction approach is clearly superior, one could test the prediction approach for a portion of the period before the alleged anticompetitive behavior, as discussed above in part D. If predicted prices before the alleged anticompetitive period are close to the actual prices, then the prediction approach presumably would be more accurate in this example. If the prediction model does not predict prices before the alleged anticompetitive acts well, one may conclude that neither the dummy variable model nor the prediction model is fully accurate. In such a situation, one would do well to also consider alternative model specifications as well as other sources of evidence (e.g., a comparison of profit margins in the before and during periods) to confirm that a range of evidence reaches the same result.

7. *Benchmark Approach*

In a benchmark analysis, the first question is whether there are any products that might provide an appropriate benchmark. As discussed in part F above, the benchmark should be (1) subject to similar supply and demand factors as SPRB coal; and (2) unaffected by alleged anticompetitive acts. This illustration will use the price of bituminous coal (BIT).[237]

237. Utilities also purchase this coal to generate electricity, and so BIT's prices likely reflect many of the same cost and demand factors. There is no allegation of the conspiracy extending to BIT. Moreover, most utilities cannot substitute SPRB and BIT, at least in the short run, so the alleged

An econometric analysis using the benchmark approach can be implemented in several ways. To keep the illustration relatively simple, the impact of the alleged conspiracy is tested using the dummy variable model discussed above. Most of the supply and demand factors used in the before-during approaches are replaced by the benchmark BIT price, since it should be capturing those effects.[238] The model includes the same dummy variable, as well as three variables to control for the quality of SPRB coal.[239]

Table 4: Regression Results for Benchmark Approach

Dependent Variable:	Average Price of SPRB Coal
R-squared:	0.7191
Adj R-squared:	0.7093
Number of Obs:	149

	Coefficient	Newey-West Standard Error	t-Statistic	P>|t|
Dummy variable for the alleged collusion period	1.1527	0.1867	6.1700	0.0000
Btu content of SPRB coal	-0.0005	0.0010	-0.4800	0.6300
Ash percentage of the SPRB coal	-1.1329	0.6171	-1.8400	0.0680
Pounds of sulfur per million Btu	-1.1130	1.3201	-0.8400	0.4010
Average tons of SPRB coal produced per person day	0.0063	0.0028	2.2800	0.0240
Constant	15.7235	9.7266	1.6200	0.1080

As can be seen in Table 4, the BIT price is of the expected sign (positive) and is statistically significant. The results show a statistically significant impact on SPRB prices in the period during the alleged conspiracy, again consistent with the alleged conspiracy increasing price.

---

conspiracy should have no direct effect on the price of BIT. In *Federal Trade Commission v. Arch Coal, Inc.*, the court found "the relevant product market is no broader and no narrower than SPRB coal." FTC v. Arch Coal, Inc., 329 F. Supp. 2d 109, 123 (2004).

238. The supply and demand factors could still be included in the model, and if they affect the benchmark differently than the price in question, including them will generally improve the model.

239. The estimating equation is: $SPRB\_price_t = \alpha + \beta_1 \cdot Btu_t + \beta_2 \cdot Ash_t + \beta_3 \cdot lbsSO2_t + \beta_4 \cdot BIT\_price_t + \gamma \cdot dummy_t + \varepsilon_t$, where BIT_price_t is the price of BIT coal in period t. The results of this regression are shown in Table 4.

This specification of the benchmark dummy variable model estimates a \$1.15/ton (or about 21 percent) increase in the SPRB coal price, which is somewhat smaller than the reduced form estimates in the before-during models using either the dummy variable or prediction approaches. Although BIT should not have been directly affected by the alleged anticompetitive act related to SPRB coal, it is possible that the customers who bought BIT faced substantially different supply and demand conditions than the plants purchasing the SPRB coal. In deciding whether the benchmark approach yields a more accurate quantification of the alleged anticompetitive behavior than the dummy variable or prediction approaches, one would need to analyze in more depth the comparability of the exogenous influences on BIT and SPRB to determine how good a benchmark BIT is for SPRB. If there are substantial questions about how comparable BIT is to SPRB, then the other approaches likely would yield more accurate estimates.

J. Additional Example: An Application of Structural Modeling

Although the before-during (and related) methods, as well as the benchmark method, can be effective methods for estimating damages in many circumstances, they require strong assumptions to produce reliable estimates of antitrust damages. In particular, as described above, assumptions regarding the stability of the economic relationship between the outcome being measured and the explanatory variables are critical for the reliable application of these methods. For these models to yield reliable predictions of economic outcomes in the absence of the alleged anticompetitive behavior (and thus reliable damage estimates), it must be true that the same economic model that applies in the before (or after) period or in the benchmark market would also have applied in the affected market in the but-for during period. In situations where an industry is undergoing large changes over time (e.g., many high technology industries) or where the benchmark market is different from the affected market in important ways, this critical stability assumption is less likely to hold.

If stability cannot clearly be established, other methods should be considered. As noted above, such methods should generally include the use of non-econometric analyses, such as an evaluation of whether there were clear breaks in the trend of prices at the start or end of the alleged anticompetitive behavior, a comparison of price-cost margins, and so on.

However, in some cases, more complicated structural econometric models may also be useful. For example by relying directly on economic

theory, structural econometric methods may enable one to separate collusive from non-collusive outcomes using data entirely from within the affected market in the during period. In this way, questionable assumptions about the stability of the appropriate model in the affected period and in benchmark periods/markets can be avoided.

An excellent example of this approach—although not in the context of litigation—is found in a recent published paper by economist Aviv Nevo on the "ready-to-eat" cereal industry, which is known to have high price-cost margins.[240] A relevant antitrust question is whether these high price-cost margins reflect collusive behavior, or whether they reflect non-collusive decisions made by firms based on customer loyalty to specific brands and the fact that firms control several such brands, with the risk of cross-brand cannibalization reducing their incentives to cut prices.[241] Nevo's paper seeks to answer this question without use of a non-collusive benchmark period or industry.

Nevo's approach is to use a structural model to predict price-cost margins in the presence or absence of collusion and then to compare these price-cost margins to those actually observed. To do so, he estimates an explicitly specified structural model of the demand for specific brands of cereal. Using a "discrete-choice" model, such as that found in the seminal work of Berry, Levinsohn, and Pakes,[242] Nevo's model describes the demand for 25 brands of cereal as a function of the prices and characteristics of those brands. He estimates the model using supermarket scanner data from roughly 65 U.S. cities over a period of 20 quarters.

The central econometric challenge that Nevo faces is that cereal prices are likely endogenous, meaning that they are correlated with unobservable factors (i.e., in the error term) that affect cereal demand. In particular, cereals for which consumers have greater demand, for reasons that are not fully observable, likely also have higher prices reflecting their greater desirability. If one does not account for this problem in estimation, the positive correlation between price and unobserved demand factors may make it appear that consumers actually prefer higher prices, or at least may bias the negative relationship between price and demand toward zero.

---

240. Aviv Nevo, *Measuring Market Power in the Ready-to-Eat Cereal Industry*, 69 ECONOMETRICA 307-342 (2001).

241. Nevo's focus is on the possibility of "tacit collusion" among cereal manufacturers, not explicit, unlawful price-fixing. However, this distinction is immaterial to the methods discussed here, which could also be used to detect explicit collusion.

242. *BLP, supra* note 104.

Nevo overcomes this problem using the instrumental variables techniques described above. As an instrument for the price of a given brand of cereal in a given city-quarter combination, Nevo uses the prices of that same brand of cereal in other cities in the same quarter. For this to be a valid instrument, Nevo must assume that the city-quarter specific demand "shocks" are uncorrelated across cities, while there are common cost shocks affecting multiple cities in a given quarter. In this way, prices in other cities are correlated with prices in a given city (due to common cost factors), but *not* correlated with the error term in the demand equation (since demand shocks are not correlated), making these prices in other cities valid instruments. Recognizing that these assumptions are restrictive (as is often the case in structural models), Nevo also considers the use of direct measures of city-specific marginal costs as instruments.

With the estimates of demand in place, Nevo then uses an explicit model of profit-maximizing price setting, under either separate competition by individual firms or collusion, to derive predicted price-cost margins. He then compares these predicted price-cost margins to actual measures of price-cost margins. He concludes that the ready-to-eat cereal industry is *not* characterized by collusion, as the price-cost margins are lower than those predicted by the collusive model and instead consistent with a model of individual price setting by the separate firms.

This example illustrates two key features of structural modeling. First, that by relying on predictions of economic theory as additional information, structural models may allow identification of collusive (or other) effects that cannot be identified purely from the data alone. Second, that making use of such predictions from economic theory—and especially applying those predictions to available data—often requires strong assumptions. As such, structural methods may be quite useful, but should be carefully tested and often paired with other, simpler methods to make sure the results are robust to a variety of approaches.